Network Working Group
Internet Draft
Category: Standards Track
Expiration Date: April 2005

George Swallow
Cisco Systems, Inc.

Kireeti Kompella
Juniper Networks, Inc.

Dan Tappan
Cisco Systems, Inc.

October 2004

# Label Switching Router Self-Test

draft-ietf-mpls-lsr-self-test-03.txt

Status of this Memo



Copyright Notice



Abstract

This document defines a means of self test for a Label-Switching Router (LSR) to verify that its dataplane is functioning for certain key Multi-Protocol Label Switching (MPLS) applications including unicast forwarding based on LDP [LDP] and traffic engineering tunnels based on [RSVP-TE]. A new Loopback FEC type is defined to allow an upstream neighbor to assist in the testing at very low cost. MPLS Echo Request and MPLS Echo Reply messages [LSP-Ping] are extended to do the actual probing.

Contents

1 Introduction ..... 3
1.1 Conventions ..... 3
2 Loopback FEC ..... 4
2.1 Loopback FEC Element ..... 4
2.2 LDP Procedures ..... 5
3 Data Plane Self Test ..... 5
3.1 Data Plane Verification Request / Reply Messages ..... 6
3.2 Reply-To Object ..... 8
3.2.1 IPv4 Reply-To Object ..... 8
3.2.2 IPv6 Reply-To Object ..... 8
3.3 Sending procedures ..... 9
3.4 Receiving procedures ..... 10
3.5 Upstream Neighbor Verification ..... 11
4 Security Considerations ..... 11
5 IANA Considerations ..... 12
6 Acknowledgments ..... 12
7 References ..... 12
7.1 Normative References ..... 12
7.2 Informative References ..... 12
8 Authors' Addresses ..... 13
9 Full Copyright and Intellectual Property Statements ..... 13

## 1. Introduction

This document defines a means of self test for a Label-Switching Router (LSR) to verify that its dataplane is functioning for certain key Multi-Protocol Label Switching (MPLS) applications including uni-cast forwarding based on LDP [LDP] and traffic engineering tunnels based on [RSVP-TE]. MPLS Echo Request and MPLS Echo Reply messages [LSP-Ping] messages are extended to do the actual probing. The pings are sent to an upstream neighbor, looped back through the LSR under test and intercepted, by means of TTL expiration by a downstream neighbor. Extensions to LSP-Ping [LSP-Ping] are defined to allow the down stream neighbor to report the test results.

In order to minimize the load on upstream LSRs a new loopback FEC is defined. Receipt of a packet labeled with a loopback label will cause the advertising LSR to pop the label off the label stack and send the packet out the advertised interface.

Note that use of a loopback allows an LSR to test label entries for which the LSR is not currently some neighbor's next hop. In this way label entries can be verified prior to the occurrence of a routing change.

Some routing protocls, most notably OSPF have no means of exchanging the "Link Local Identifiers" used to identify unnumbered links and components of bundled links. These test procedures can be used to associate the neighbor's interfaces with the probing LSRs interfaces. This is achieved by simply having the TTL of the MPLS Ping expire one hop sooner, i.e. at the testing LSR itself.

### 1.1. Conventions

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 [KEYWORDS].

## 2. Loopback FEC

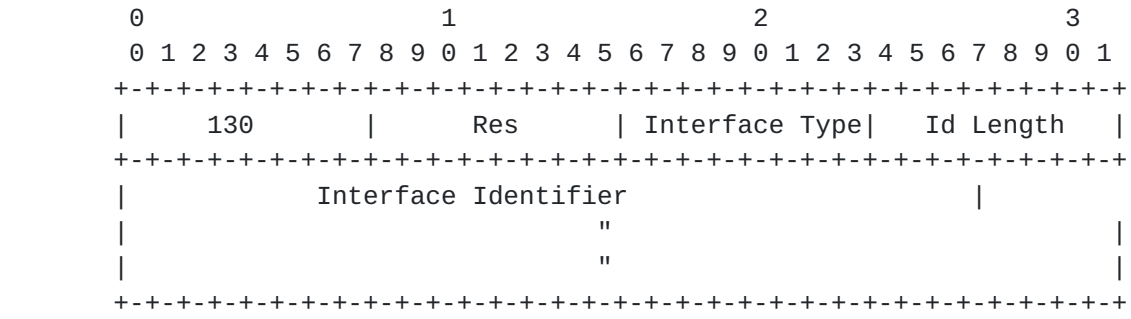
The Loopback FEC type is defined to enable an upstream neighbor to assist in LSR self-testing at very low cost. This FEC causes the loopback to occur in the dataplane without control plane involvement beyond the initial LDP exchange and dataplane setup.

An LSR uses the Loopback FEC to selectively advertise loopback labels to its neighbor LSRs. Each loopback label is bound to a particular interface. For multi-access links, a unique label for each neighbor is required, since the link-level address is derived from the label lookup. When an MPLS packet with its top label set to a loopback label is received from an interface over which that label was advertised, the loopback label is popped and the packet is sent on the interface to which the loopback label was bound.

TTL treatment for loopback labels follows the Uniform model. I.e. the TTL carried in the loopback label is decremented and copied to the exposed label or IP header as the case may be.

### 2.1. Loopback FEC Element

FEC element type 130 is used. The FEC element is encoded as follows: (note: 130 is provisionally assigned, the actual value will be assigned by IANA.)




Reserved (Res)

Must be set to zero on transmission and ignored on receipt.

Interface Type

| # | Type | Interface Identifier |
|---|---|---|
| 0 | Unnumbered | A 32 bit Link Identifier as defined in [RFC3477] |
| 1 | IPv4 Numbered | IPv4 Address |
| 2 | IPv6 Numbered | IPv6 Address |

Identifier Length

Length of the interface identifier in octets. The length is 4 bytes for Unnumbered and IPv4, 16 bytes for IPv6.

Address

An identifier encoded according to the Identifier Type field.

### 2.2. LDP Procedures

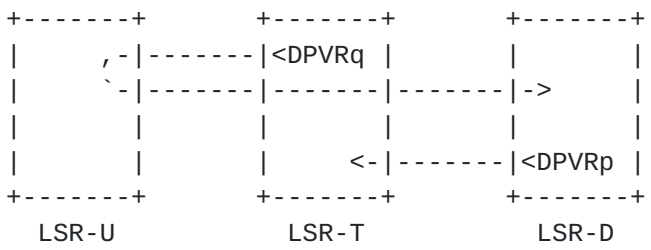
It is RECOMMENDED that loopback labels only be distributed in response to a Label Request message, irrespective of the label advertisement mode of the LDP session. However it is recognized that in certain cases such as OSPF with unnumbered links, the upstream LSR may not have sufficiently detailed information of the neighbor's link identifier to form the request. In these cases, the downstream LSR will need to be configured to make unsolicited advertisements.

## 3. Data Plane Self Test

A self test operation involves three LSRs, the LSR doing the test, an upstream neighbor and a downstream neighbor. We refer to these as LSRs T, U, and D respectively. In order to minimize the processing load on LSR-D, two new LSP Ping messages are defined, called the MPLS Data Plane Verification Request and the MPLS Data Plane Verification Reply. These messages are used to allow LSR-T to obtain the label stack, address and interface information of LSR-D.

If FEC verification is required, the MPLS Echo Request and Reply messages are used.

The packet flow is shown below. Although the figure shows LSR-D adjacent to LSR-T it may in some cases be an arbitrary number of hops away.




DPVRq: MPLS Data Plane Verification Request
DPVRp: MPLS Data Plane Verification Reply

Figure 1: Self Test Message Flow

In order to perform a test on an incoming label stack, LSR-T forms an MPLS Data Plane Verification Request. LSR-T prepends the packet with the incoming label stack being tested and the loopback label received from LSR-U. The TTL values are set such that they will expire at LSR-D. LSR-T then forwards the packet to LSR-U.

LSR-U receives the packet and performs normal MPLS forwarding. That is, the loopback label is popped, the TTL is decremented and propagated (in this case) to the exposed label.

LSR-T receives the packet and performs normal MPLS forwarding. If everything is functioning as expected this will cause the packet to arrive at LSR-D with a TTL of 1.

In this example, we assume that all is working properly. The TTL expires at LSR-D causing it to receive the packet. LSR-D notes the the interface and the label stack on which the packet was received and records these in an Interface and Label Stack TLV. This Object is sent to LSR-T in an MPLS Data Plane Verification Reply message.

## 3.1. Data Plane Verification Request / Reply Messages

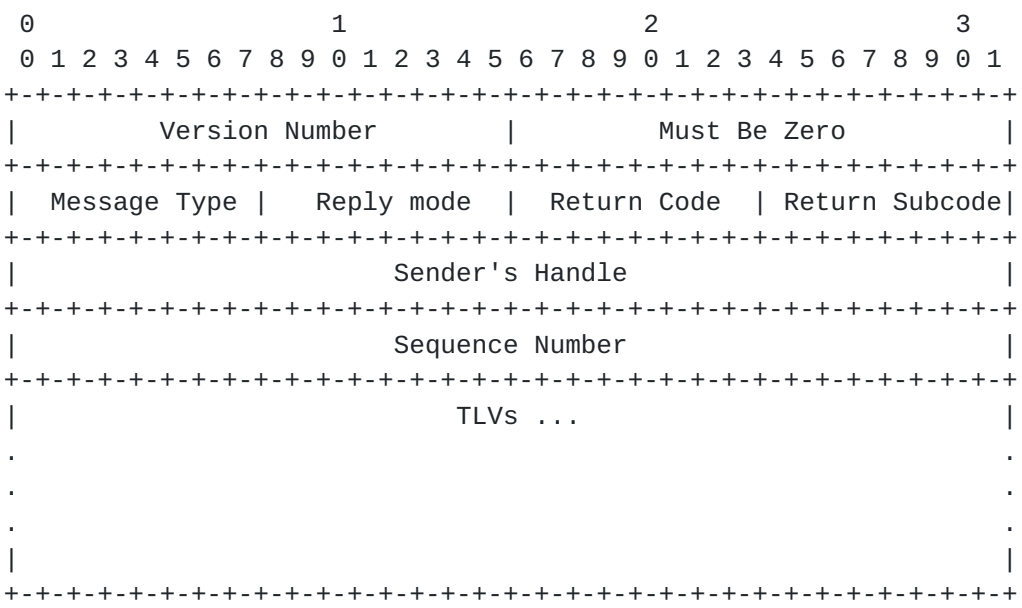
Two new LSP Ping messages are defined for LSR self test. The purpose of the new messages is three fold. First the timestamps are removed to minimize processing. Second the message type allows simple recognition that minimal processing is necessary to service this request. Third, the Verification Request message itself conveys the the request, thus a Verification Request message with no Objects is both legal and normal.

The definitions of all fields in the messages are identical to those found in [LSP-PING].

The new message types are: (Provisionally; to be assigned)

| Type | Message |
|---|---|
| 3 | MPLS Data Plane Verification Request |
| 4 | MPLS Data Plane Verification Reply |

The messages have the following format:




The MPLS Data Plane Verification Request message MAY contain the following objects:

| Type # | Object |
|---|---|
| 3 | Pad |
| 5 | Vendor Enterprise Code |
| 9 (tba) | IPv4 Reply-to Object |
| 10 (tba) | IPv6 Reply-to Object |

The MPLS Data Plane Verification Reply message MAY contain the following objects:

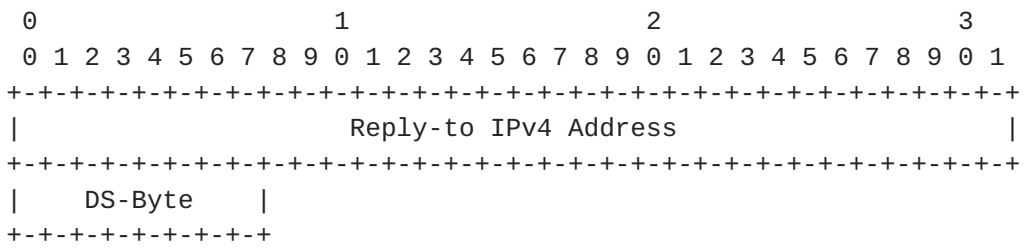
| Type # | Object |
|---|---|
| 3 | Pad |
| 4 | Error Code |
| 5 | Vendor Enterprise Code |
| 7 (tba) | IPv4 Interface and Label Stack Object |
| 8 (tba) | IPv6 Interface and Label Stack Object |

### 3.2. Reply-To Object

In order to perform detailed diagnostics of a particular failing flow in the face of ECMP, it is useful to be able to use the exact source and destination addresses of that flow. The Reply-To Object is an optional TLV in a MPLS Data Plane Verification Request message. The Object has two formats, type 9 for IPv4 and type 10 for IPv6 (to be assigned by IANA).

#### 3.2.1. IPv4 Reply-To Object

The length of an IPv4 Reply-To Object is 5 octets; the value field has the following format:




Reply-to IPv4 Address

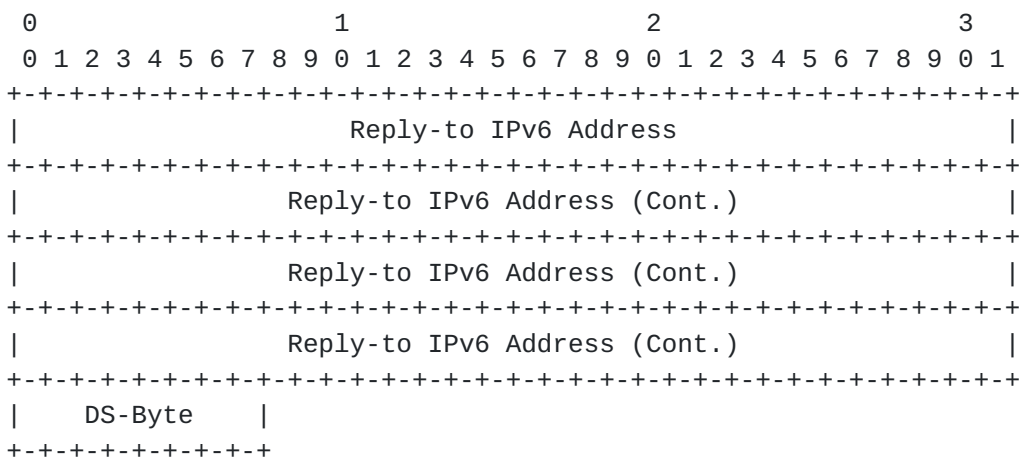
The address to which the MPLS Data Plane Verification Reply message is to be sent.

DS-Byte

The DS-Byte to be used in the MPLS Data Plane Verification Reply packet.

#### 3.2.2. IPv6 Reply-To Object

The length of an IPv6 Reply-To Object is 17 octets; the value field has the following format:




Reply-to IPv6 Address

The address to which the MPLS Data Plane Verification Reply message is to be sent.

DS-Byte

The DS-Byte to be used in the MPLS Data Plane Verification Reply packet.

## 3.3. Sending procedures

In order to perform a test on an incoming label stack, an LSR first determines the expected outgoing label stack, next hop router and next hop interface.

The LSR creates an MPLS Data Plane Verification Request message and includes a Data Plane Verification Object. Optionally a FEC Stack TLV may be included. In this case an MPLS Echo Request Message MUST be used.

In normal use, the source address is set to an address belonging to the LSR and the destination set to an address in the range of 127/8. The IP TTL SHOULD be set to 1. The incoming label stack is prepended to the packet. The TTL of these labels SHOULD be set to appropriate values - 2 for those labels which will be process by this when the packet is looped back; 1 for those labels which will be carried through. Finally the loopback label bound to the incoming interface is prepended to the packet. The TTL is set such that it will have the value of 3 on the wire.

The packet is sent to the upstream neighbor on an interface for which the loopback label is valid.

In diagnostic situations, the source and destination addresses MAY be set to any value. In this case, a Reply-to IPv4 or IPv6 Object MUST be included. The IP TTL MUST be set to 1. The TTL of labels other than the loopback label MUST be set to appropriate values - 2 for those labels which will be process by this LSR when the packet is looped back; 1 for those labels which will be carried through.

### 3.4. Receiving procedures

An LSR X that receives an MPLS Verification Request message formats a MPLS Verification Reply message. The Sender's Handle and Sequence Number are copied from the Request message.

X then parses the packet to ensure that it is a well-formed packet, and that the TLVs that are not marked "Ignore" are understood. If not, X SHOULD send an MPLS echo reply with the Return Code set to "Malformed echo request received" or "TLV not understood" (as appropriate), and the Subcode set to zero. In the latter case, the misunderstood TLVs (only) are included in the reply.

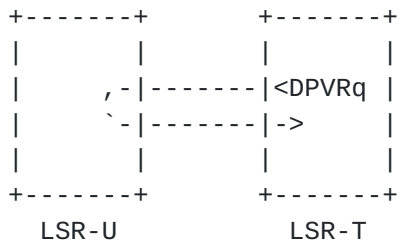
If the echo request is good, X notes the interface I over which the echo was received, and the label stack with which it came. If the MPLS echo request contained a Downstream Verification object, then X must format this information as a Downstream Verification object and include it in its MPLS echo reply message.

The source address of the Reply message MUST be an address of the replying LSR. If the request included a Reply-to IPv4 or IPv6 Object, the MPLS Data Plane Verification Reply message MUST be sent to that address. Otherwise the Reply message is sent to the source address of the Verification Request message.

An LSR MUST be capable of filtering addresses that are to be replied to. If a filter has been invoked (i.e. configured) and an address does not pass the filter, then a reply MUST NOT be sent, and the event SHOULD be logged.

### 3.5. Upstream Neighbor Verification

To verify that an upstream neighbor is properly echoing packets an LSR may send an MPLS Data Plane Verification Request packet with the TTL set so that the packet will expire upon reaching reaching itself. This procedure not only tests that the neighbor is correctly processing the loopback label, it also allow the node to verify the neighbor's interface mapping.




DPVRq: MPLS Data Plane Verification Request

Figure 2: Upstream Neighbor Verification

No TLVs need to be included in the MPLS Data Plane Verification Request. By noting the Sender's Handle and Sequence Number, as well as the loopback label, LSR-T is able to detect that a) the packet was looped, and b) determine (or verify) the interface on which the packet was received.

## 4. Security Considerations

Were loopback labels widely known, they might be subject to abuse. It is therefore RECOMMENDED that loopback labels only be shared between trusted neighbors. Further, if the loopback labels are drawn from the Global Label Space, or any other label space shared across multiple LDP sessions, it is RECOMMENDED that all loopback labels be filtered from a session except those labels pertaining to interfaces directly connected to the neighbor participating in that session.

## 5. IANA Considerations

TBD

## 6. Acknowledgments

The authors would like to thank Vanson Lim, Tom Nadeau, and Bob Thomas for their comments and suggestions.

## 7. References

### 7.1. Normative References

[RFC3036] Andersson, L. et al., "LDP Specification", January 2001.

[LSP-Ping] Bonica, R. et al., "Detecting MPLS Data Plane Liveness", work-in-progress.

[RFC3477] Kompella, K. & Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", January 2003.

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

### 7.2. Informative References

[RSVP-TE] Awduche, D., et al, "RSVP-TE: Extensions to RSVP for LSP tunnels", RFC 3209, December 2001.

## 8. Authors' Addresses

Kireeti Kompella
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
Email: kireeti@juniper.net

George Swallow
Cisco Systems, Inc.
1414 Massachusetts Ave
Boxborough, MA 01719

Email: swallow@cisco.com

Dan Tappan
Cisco Systems, Inc.
1414 Massachusetts Ave
Boxborough, MA 01719

Email: tappan@cisco.com

## 9. Full Copyright and Intellectual Property Statements



Intellectual Property

made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

## Acknowledgement

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