Network Working Group
Internet Draft
Category: Standards Track
Expiration Date: April 2004

George Swallow
Cisco Systems, Inc.

Kireeti Kompella
Juniper Networks, Inc.

Dan Tappan
Cisco Systems, Inc.

October 2003

Label Switching Router Self-Test

draft-ietf-mpls-lsr-self-test-01.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 actually probing.

Contents

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 unicast 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 actually 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 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.)

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |     130       |       Res     | Interface Type|   Id Length   |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                     Interface Identifier                      |
   |                               "                               |
   |                               "                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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 neighbors 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 LSRD, 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 LSRT to obtain the label stack and address and interface information of LSRD.

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

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

```
                 +-------+      +-------+      +-------+
                 |    ,-|-------|<DPVRq|       |       |
                 |    `-|-------|------|-------|->     |
                 |      |       |    <-|-------|<DPVRp|
                 +-------+      +-------+      +-------+
                   LSRU           LSRT           LSRD

               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, LSRT forms an MPLS Data Plane Verification Request. Included in that is a Data Plane Verification Object which requests that the interface and label stack seen by LSRD be returned. Optionally a FEC Stack TLV may be included to verify LSRD's labels are mapped to the expected FECs. In this case a MPLS Echo Request Message MUST be used.

LSRT prepends the packet with the incoming label stack and the loopback label received from LSRU. The TTL values are set such that they will expire at LSRD. LSRT then forwards the packet to LSRU.

LSRU 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.

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

LSRD receives the packet. It verifies that the packet was received on the interface and with the label stack contained in the Next Hop Verification TLV. If a FEC Stack TLV is also included, it verifies that the labels in the stack are mapped to those FECs. The results are recorded in an MPLS Data Plane Verification Reply message and sent to LSRT.

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 two 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. The definitions of all fields 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:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Version Number        |         Must Be Zero          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Message Type |   Reply mode  |  Return Code  | Return Subcode|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sender's Handle                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            TLVs ...                           |
.                                                               .
.                                                               .
.                                                               .
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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

Type #	Object
3	Pad
5	Vendor Enterprise Code
7 (tba)	IPv4 Downstream Verification Object
8 (tba)	IPv6 Downstream Verification Object
9 (tba)	IPv4 Reply-to Object
10 (tba)	IPv6 Reply-to Object

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

Type #	Object
3	Pad
4	Error Code
5	Vendor Enterprise Code
7 (tba)	IPv4 Downstream Verification Object
8 (tba)	IPv6 Downstream Verification Object
9 (tba)	IPv4 Reply-to Object
10 (tba)	IPv6 Reply-to Object

3.2. Next Hop Verification Object

The Down Stream Verification Object is an optional TLV in an echo request. Only one such object may appear. It's presence signifies a request that Next Hop Verification Object be included in the echo reply. The purpose of the object is to allow the upstream router to obtain the exact interface and label stack information as it appears at the replying LSR. It has two formats, type 7 for IPv4 and type 8 for IPv6 (to be assigned by IANA).

3.2.1. IPv4 Next Hop Verification Object

In an echo request message the Length is always 12. In an echo reply message the length is 16 + 4*N octets, N is the number of Downstream Labels. The value field of a Next Hop Verification TLV has the following 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Flags     | Address Type  |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Downstream IPv4 Address                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Downstream Interface Address                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                                                               .
.                          Label Stack                          .
.                                                               .
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Flags

Two flags are defined as shown. All other flags are reserved and MUST be set to zero.

```
+-+-+-+-+-+-+-+-+
|0|0|0|0|0|0|V|R|
+-+-+-+-+-+-+-+-+
```

The R flag can only be set in an echo request message. It requests that the receiving router verify that the Downstream IPv4 Address is an address belonging to this router.

The V flag can only be set in an echo reply message. The flag is set as a positive verification response to a received R flag. If the R flag was not set in the echo request this Flag MUST be set to zero.

Address Type

The Address Type indicates if the interface is numbered or unnumbered and is set to one of the following values:

Address Type	Value
No Address	0
IPv4	1
Unnumbered	2

The value 0, "No Address" is only valid in a Verify Request message.

Downstream IPv4 Address

If the address type is 'No Address', the address field MUST be set to zero and ignored on receipt.

If the address type is 'IPv4', the address field MUST either be set to the downstream LSR's Router ID or the downstream LSR's interface address.

If the address type is 'unnumbered', the address field MUST be set to the downstream LSR's Router ID.

Downstream Interface Address

If the address type is 'IPv4', the interface address field MUST MUST be set to the downstream LSR's interface address.

If the address type is 'unnumbered', interface address field MUST be set to the index assigned by the downstream LSR to the interface.

Reserved

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

Label Stack

The label stack of the received echo request message. If any TTL values have been changed by this router, they SHOULD be set restored.

3.2.2. IPv6 Next Hop Verification Object

In an echo request message the Length is always 24. In an echo reply message the length is 40 + 4*N octets, N is the number of Downstream Labels. The value field of a Next Hop Verification TLV has the following 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Flags     | Address Type  |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Downstream IPv6 Address                    |
|                Downstream IPv6 Address (Cont.)                |
|                Downstream IPv6 Address (Cont.)                |
|                Downstream IPv6 Address (Cont.)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Downstream Interface Address                 |
|              Downstream Interface Address (Cont.)             |
|              Downstream Interface Address (Cont.)             |
|              Downstream Interface Address (Cont.)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                                                               .
.                          Label Stack                          .
.                                                               .
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Flags

Two flags are defined as shown. All other flags are reserved and MUST be set to zero.

```
+-+-+-+-+-+-+-+-+
|0|0|0|0|0|0|V|R|
+-+-+-+-+-+-+-+-+
```

The R flag can only be set in an echo request message. It requests that the receiving router verify that the Downstream IPv6 Address is an address belonging to this router.

The V flag can only be set in an echo reply message. The flag is set as a positive verification response to a received R flag. If the R flag was not set in the echo request this Flag MUST be set to zero.

Address Type

The Address Type indicates if the interface is numbered or unnumbered and is set to one of the following values:

Address Type	Value
No Address	0
IPv6	1
Unnumbered	2

The value 0, "No Address" is only valid in a Verify Request message.

Downstream IPv6 Address

If the address type is 'No Address', the address field MUST be set to zero and ignored on receipt.

If the address type is 'IPv6', the address field MUST either be set to the downstream LSR's Router ID or the downstream LSR's interface address.

If the address type is 'unnumbered', the address field MUST be set to the downstream LSR's Router ID.

Downstream Interface Address

If the address type is 'IPv6', the interface address field MUST MUST be set to the downstream LSR's interface address.

If the address type is 'unnumbered', first four octets of interface address field MUST be set to the index assigned by the downstream LSR to the interface. The remaining 12 octects MUST be set to zero.

Reserved

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

Label Stack

The label stack of the received echo request message. If any TTL values have been changed by this router, they SHOULD be set restored.

3.3. Reply-To Object

In order to perform detailed diagnostics of a particular failing flow in the face of EMCP, 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 data plane verify request. The Object has two formats, type 9 for IPv4 and type 10 for IPv6 (to be assigned by IANA).

3.3.1. IPv4 Reply-To Object

The length of an IPv4 Reply-To Objectis 5 octets; the value field has the following 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Reply-to IPv4 Address                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    DS-Byte    |
+-+-+-+-+-+-+-+-+
```

Reply-to IPv4 Address

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

DS-Byte

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

3.3.2. IPv6 Reply-To Object

The length of an IPv6 Reply-To Objectis 17 octets; the value field has the following 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Reply-to IPv6 Address                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Reply-to IPv6 Address (Cont.)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Reply-to IPv6 Address (Cont.)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Reply-to IPv6 Address (Cont.)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    DS-Byte    |
+-+-+-+-+-+-+-+-+
```

Reply-to IPv6 Address

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

DS-Byte

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

3.3.3. 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 packet and includes a Data Plane Verification Object. 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 itself 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

The processing rules of the Downstream Data Plane Verification Object are followed. If the request included a Reply-to IPv4 or IPv6 Object, the MPLS Data Plane Verification Reply message MUST be sent to that address. The source address MUST be an address of the replying LSR.

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
           |    ,-|-------|<DPVRq|           Verification Request
           |    `-|-------|->    |
           |      |       |      |
           +-------+       +-------+
             LSRU           LSRT
```

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, LSRT 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 label 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