Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 26, 2012

Q. Zhao
E. Chen
Huawei Technology

October 24, 2011

# Protection Mechanisms for Label Distribution Protocol P2MP/MP2MP Label Switched Paths

**draft-zhao-mpls-mldp-protections-00.txt**

## Abstract

Service providers continue to deploy real-time multicast applications using Multicast LDP (mLDP) across MPLS networks. There is a clear need to protect these real-time applications and to provide the shortest switching times in the event of failure. This document outlines the requirements, describes the protection mechanisms available, and where neccessary proposes extensions to facilitate mLDP P2MP and MP2MP LSP protection within an MPLS network.

## Status of this Memo



## Copyright Notice

Table of Contents

## 1. Terminology

For a clear narrative, this section gives a general conceptional overview of the terms.

- PLR: The node where the traffic is logically redirected onto the preset backup path is called Point of Local Repair.

- MP: The node where the backup path merges with the primary path is called Merge Point.

- FD: The node that detects the failure on primary path, and then triggers the action(s) for traffic protection is called Failure Detector. Either traffic sender or receiver can be the FD, depending on which protection mode are deployed. More details are described in later sections of this document.

- SP: The node where the traffic is physically switched/duplicated onto the backup path is called Switchover Point. In multicast cases, PLR and SP can be two different nodes. More details are described in later sections of this document.

## 2. Requirement Language

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].

## 3. Introduction

In order to meet user demands, operators and service providers continue to deploy multicast applications using mLDP across MPLS networks. In certain scenarios, traditional IGP-mLDP convergence mechanisms fail to meet protection switching times required to minimise, or negate entirely, application interruptions for real-time applications, including stock trading, on-line games, and multimedia teleconferencing.

Current best practice for protecting services, and higher applications includes the pre-computation and establishment of a backup path, this can decrease the convergence time efficiently. Once a failure has been detected on the primary path, the traffic should be transmitted across the back-up path.

However, two major challenges exist with the aforementioned solution. The first is how to build an absolutely disjointed backup path for

each node in a multicast tree; the second is how to balance between convergence time and resource consumption.

This document provides several ways to setup the backup path for mLDP LSP, including local protection, territorial protection, and end-to-end protection. The goal is to build a reliable umbrella to against traffic black hole. How to detect failure is outside the scope of this document.

More and more users are apt to deploy multicast applications on MPLS mLDP network. In some scenarios, traditional IGP-mLDP convergence is hard to meet the requirements of those real-time applications, such as stock business, on-line game, and multimedia teleconference.

The industry has reached a consensus that setting up a backup path previously can decrease the convergence time efficiently. No matter how the above-mentioned backup path was established, once the failure is detected, the traffic should be transmitted at that path as soon as possible. Even so, there are still two major challenges left for us, one is how to build an absolutely disjointed backup path for each node in a multicast tree; the other is how to balance between convergence time and resource consumption.

It is getting urgent to find the ideal protection mechanism(s) to improve the convergence time, and at the meantime minimize the side-effects, such as bandwidth wastage.

For a primary LDP P2MP/MP2MP LSP, there are several ways to set up its backup path. It can use RSVP-TE P2P tunnel as a logical out-going interface, consequently utilize the mature high availability technologies of RSVP-TE. Or, it can make use of LDP P2P backup LSP as a packet encapsulation, so that the complex configuration of P2P RSVP-TE can be skipped. Or, it can build its own P2MP/MP2MP backup LSP according to IGP's loop-free alternative route, thus avoid double label stack. Other than these, it can also build a totally disjointed LSP in another topology, accordingly take advantage of the real end-to-end protection.

When the backup path is present, there are two options for packet forwarding and switchover. If the traffic sender feeds the stream on both paths, and the traffic receiver drops packet on backup path, the switchover will be very quick once the failure is detected, because the whole switchover action is a local behavior on traffic receiver. The disadvantage of this manner is that traffic will be duplicated on both paths, and consume double bandwidth. Contrastively, if the traffic sender feeds stream only on the primary path, the resource wastage can be waived. Cooperation is needed in this manner, so there will be some protocol extensions. But if the performance can

be equal or better than the previous option, it is reasonable to choose the second one.

This document describes several methods to setup and switch paths for options to setup the backup LDP P2MP/MP2MP LSP. mLDP LSPs, including local protection, territorial protection, and end-to-end protection. The goal is to identify strengths, weaknesses and gaps, in order to build a reliable set of tools to shield against traffic black holes that would severely impact real-time applications, in the event of primary path failure.

## 3.1. Requirements

A number of requirements have been identified that allow the optimal set of mechanisms to developed. These currently include:

- o Computation of a disjointed (link and node) backup path within the multicast tree;

- o Minimisation of protection convergence time;

- o Optimisation of bandwitdth usage.

## 3.2. Scope

The method to detect failure is outside the scope of this document. Also this document does not provide any authorization mechanism for controlling the set of LSRs that may attempt to join a mLDP protection session.

# 4. Local protection using P2P LSP

By encapsulating mLDP packets within an P2P TE tunnel or P2P LDP backup LSP, the LDP P2MP/MP2MP LSP can be protected by the P2P protection mechanisms. However, this protection mechanism is not capable of detecting, and recovering, if the failure occurs on the destination node of the P2P backup LSP. Thus, this section provides a unified method to protect both node and link with P2P backup LSP.

```
                             +------------+ Point of Local Repair/
                             |     R1     | Switchover Point
                             +------------+ (Upstream LSR)
                                /       \
                               /         \
                            10/           \20
                             /             \
                            /               \
                           /                 \
                     +----------+        +-----------+
                     |    R2    |        |     R3    |
                     +----------+        +-----------+
                       |       \                  |
                       |        \                 |
                       |         \                |
                     10|        10\             20|
                       |           \              |
                       |            \             |
                       |             \            |
                       |              \           |
                       |               \          |
                       |                \         |
                     +-----------+  10  +-----------+ Merge Point
                     |    R4     |------|    R5     | (Downstream LSR)
                     +-----------+      +-----------+
```

mLDP Local Protection Example

Figure 1

In Figure 1 (mLDP Local Protection Example) above, the preferential path from R1 to R4/R5 is through R2, and the secondary path is through R3.  In this case, the mLDP LSP will be established according to the IGP preferential path as R1--R2--R4/R5.

It is the responsibility of R2 to inform R1 of its downstream LSRs (in this example R4 and R5) and the respective labels (L4 and L5). Once the link between R1 and R2 fails, or R2 node fails, R1 will duplicate the traffic to R4 and R5, with inner label as L4/L5, and outer label as the P2P backup LSP R1--R3--R5--R4 and R1--R3--R5.

Finally, the previous forwarding states will be removed after R4 and R5 finish their Make-Before-Break (MBB) procedure.

## 4.1. Signaling procedures for local protection

Continuing to use Figure 1 (mLDP Local Protection Example), R2 sends a notification message to R1 to inform the node that R2 has two downstream nodes, R4 and R5 with forwarding labels L4 and L5 respectively.

When R1 sees R2 node going down, it takes mLDP packets as it would send them to R4 and R5 through R2 and sends them into the two P2P backup tunnels:

- P2P tunnel R1--R3--R5--R4, using inner label L4.

- P2P tunnel R1--R3--R5, using inner label L5.

So that R4/R5 will receive same packets as from the interface between R2 and R4/R5, just from different interface.

At the same time, R1 sends notifications with MBB request status code to R4 and R5. So that after R4 and R5 are done with MBB, they will send the notifications to R3 with MBB done status code. And then R3 will remove the old forwarding state which is being protected by the P2P backup tunnels.

## 4.2. Protocol extensions for local protection

A new type of LDP MP Status Value Element is introduced, for notifying downstream LSRs and respective labels. It is encoded as follows:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|mLDP P2P Type=2|      Length                   |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Downstream Element 1                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Downstream Element N                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

mLDP P2P Encapsulation Status Code

Figure 2

The Downstream Element is encoded as follows:

```
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                   Downstream Label    |                       |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                   Downstream LSR-ID                           |
    +                               +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Downstream Element in mLDP P2P Encapsulation Status Code

Figure 3

## 5. Territorial protection using mLDP LSP

Making use of IGP-FRR results, LDP can build the backup mLDP LSP for territorial protection. Note that in some scenarios, such as the following example, Failure Detector and Point of Local Repair, Switchover Point and Merge Point can be different nodes.

```
                               +------------+ Point of Local Repair
                               |     R1     | (Upstream LSR)
                               +------------+
                                  /       \
                                 /         \
                                /           &
                               /             &
                              /               \
           Switchover Point \/_                \  Failure Detector
                   +----------+            +-----------+
                   |    R2    |            |     R3    |
                   +----------+            +-----------+
                       /     \                  /
                      /       \                /
                     /         \              /
                    /           \            /
                   /             \          /
                 \/_              \        /
           +----------+          +-----------+  Merge Point
           |    R4    |          |     R5    | (Downstream LSR)
           +----------+          +-----------+
```

mLDP Territorial Protection Example

Figure 4

In Figure 4 (mLDP Territorial Protection Example), normally R1 feeds traffic to R4 through R2, and feeds traffic to R5 through R3. Once the link between R1 and R3 fails, R1 will be the logical Point of Local Repair node, which feeds the traffic to R5 through backup path on R2. Because R2 is already receiving traffic, so that R1 does not need to take any action. It is responsibility of R2 to duplicate the traffic to R5, as a Switchover Point. In this case, as the Failure Detector, R3 will need to send out the notification to R2, in order to trigger the switchover procedure.

## 5.1. Signaling Procedures for Territorial Protection

Merge Point (R5) determines the primary and secondary paths according to the IGP-FRR results. Then it sends out label mapping message including an LDP MP Status TLV that carries a FRR Status Code to indicate the primary path and secondary path. At the same time, it triggers a reverse path for failure notification by sending out label request message with an LDP MP Status TLV. The reverse path is uniquely identified by root address, opaque value, and MP address.

When failure is detected by Failure Detector (R3), it will send out the failure notification, then traffic will switch to the secondary path.

When Merge Point (R5) sees the next hop to Root changed, it will advertise the new mapping, and the traffic will re-converge to the new primary path.

### 5.2. Protocol extensions for Territorial Protection

A new type of LDP MP Status Value Element is introduced, for setting up secondary mLDP LSP.  It is encoded as follows:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|mLDP FRR Type=3|      Length                   | Status code   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                    MP Node Address                            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

mLDP FRR Status Code

Figure 5

mLDP FRR Type:  Type 3 (to be assigned by IANA)

Length:  If the Address Family is IPv4, the Address Length MUST be 5; if the Address Family is IPv6, the Address Length MUST be 17.

Status code:  1 = Primary path for traffic forwarding (used in Label Mapping message)
2 = Secondary path for traffic forwarding (used in Label Mapping message)
3 = Reverse path for failure notification (used in Label Request message)
4 = Failure notification (used in Notification message)

MP Node Address:  A host address encoded according to the Address Family of this LSP.

mLDP Bandwidth Reservation Status Code Parameters

Figure 6

## 6. End-to-end protection using LDP Multiple Topology

[I-D.ietf-mpls-ldp-multi-topology] provides a mechanism to setup disjointed LSPs within different topologies. So that applications can use these redundant LSPs for end-to-end protection.

```
                             +------------+ Point of Local Repair/
                             |    Root    | Switchover Point
                             +------------+
                                /       *
                               /         *
                              /           *
                             /             *
                            /               *
                           /                 *
                   +----------+            +-----------+
                   |    R1    |            |     R2    |
                   +----------+            +-----------+
                       |   \                   *  *
                       |    \                 *   *
                       |     \               *    *
                       |      \             *     *
                       |       - - - * * * *      *
                       |           *  \           *
                       |          *    \          *
                       |         *      \         *
                       |        *        \        *
                       |       *          \       *
                     +-----------+      +-----------+
                     |  Leaf 1   |      |  Leaf 2   |
                     +-----------+      +-----------+
```

mLDP End-to-end Protection Example

Figure 7

In Figure 7 (mLDP End-to-end Protection Example), there are two separated topologies from Root node to Leaf 1 and Leaf 2. For the same FEC element, the Leaf node can trigger mLDP LSPs in each topology. Root node can setup 1:1 or 1+1 end-to-end protection, using these two mLDP LSPs.

### 6.1. Signaling Procedures for End-to-end Protection

Using Figure 7 (mLDP Local Protection Example), Leaf 1 and Leaf 2 may trigger mLDP LSPs in different topologies, sending label mapping

messages with same FEC element, different MT-ID and different label. When the Root node receives the label mapping messages from different topologies, it will set up two mLDP LSPs for application as end-to-end protection. Failure detection for the primary mLDP LSP is outside the scope of this document. But either Root node or Leaf node can be the Failure Detector.

### 6.2. Protocol extensions for End-to-end Protection

The protocol extensions required to build mLDP LSPs in different topologies is defined in [I-D.ietf-mpls-ldp-multi-topology].

## 7. Acknowledgements

We would like to thank authors of draft-ietf-mpls-mp-ldp-reqs and the authors of draft-ietf-mpls-ldp-multi-topology from which some text of this document has been inspired.

## 8. IANA Considerations

This memo includes the following requests to IANA:

- o mLDP P2P Encapsulation type for LDP MP Status Value Element.

- o mLDP FRR type for LDP MP Status Value Element.

## 9. Manageability Considerations

[Editors Note - This section requires further discussion]

### 9.1. Control of Function and Policy

### 9.2. Information and Data Models

### 9.3. Liveness Detection and Monitoring

### 9.4. Verifying Correct Operation

### 9.5. Requirements on Other Protocols and Functional Component

### 9.6. Impact on Network Operation

### 9.7. Policy Control

## 10. Security Considerations

The same security considerations apply as for the base LDP specification, as described in [RFC5036]. The protocol extensions specified in this document do not provide any authorization mechanism for controlling the set of LSRs that may attempt to join a mLDP protection session. If such authorization is desirable, additional mechanisms, outside the scope of this document, are needed.

Note that authorization policies should be implemented and/or configure at all the nodes involved .

## 11. References

### 11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", RFC 5036, October 2007.

[RFC5561] Thomas, B., Raza, K., Aggarwal, S., Aggarwal, R., and JL. Le Roux, "LDP Capabilities", RFC 5561, July 2009.

[RFC6348] Le Roux, JL. and T. Morin, "Requirements for Point-to-Multipoint Extensions to the Label Distribution Protocol", RFC 6348, September 2011.

[I-D.ietf-mpls-ldp-p2mp] Minei, I., Wijnands, I., Kompella, K., and B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", draft-ietf-mpls-ldp-p2mp-15 (work in progress), August 2011.

[I-D.ietf-mpls-ldp-multi-topology] Zhao, Q., Fang, L., Zhou, C., Li, L., So, N., and R. Torvi, "LDP Extension for Multi Topology Support", draft-ietf-mpls-ldp-multi-topology-00 (work in progress), October 2011.

### 11.2. Informative References

[RFC3468] Andersson, L. and G. Swallow, "The Multiprotocol Label Switching (MPLS) Working Group decision on MPLS signaling protocols", RFC 3468, February 2003.

## Authors' Addresses

Quintin Zhao
Huawei Technology
125 Nagog Technology Park
Acton, MA 01719
US

Email: quintin.zhao@huawei.com

Emily Chen
Huawei Technology
2330 Central Expressway
Santa Clara, CA 95050
US

Email: emily.chenying@huawei.com