

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 1, 2014

Quintin Zhao
Tao Chou
Huawei Technology
Boris Zhang
Telus Communications
Emily Chen
June 30, 2013

**P2MP Based mLDP Node Protection Mechanisms for mLDP LSP
draft-zhao-mpls-mlbp-protections-04.txt**

Abstract

This document outlines the procedures and protocol extensions for the protection of mLDP nodes within Multi-Protocol Label Switching (MPLS) networks using P2MP-based backup LSPs.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 1, 2014 .

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as

described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1.	Terminology	3
2.	Requirement Language	3
3.	Introduction	3
3.1.	Requirements	5
3.2.	Scope	5
4.	mLDP Node Protection Example	5
5.	Signaling Procedures for P2MP Based Node Protection	7
5.1.	The Computation of the Backup Path	7
5.2.	The Procedures for Merge Point, Protected Node, Pn Node and PLR Node	8
5.2.1.	Merge Point's Procedures	8
5.2.2.	Protected Node's Procedures	8
5.2.3.	Pn Node's Procedures	8
5.2.4.	PLR Node's Procedures	9
5.3.	PLR Switching Over Considerations	9
5.4.	The Procedures after the Reconvergence	10
5.5.	Considerations for MP2MP LSP Node Protection	10
5.5.1.	MP Node Procedure	11
5.5.2.	PLR Node Procedure	12
5.5.3.	Switchover Procedure	12
5.6.	Protocol Extensions for mLDP Based Node Protection	12
5.6.1.	mLDP Based MP Protection Capability Parameter TLV	12
5.6.2.	mLDP Based MP Node Protection Status Elements	13
5.6.3.	mLDP Backup FEC Element Encoding	13
5.7.	IANA Considerations	15
6.	Security Considerations	15
7.	Acknowledgements	16
8.	References	16
8.1.	Normative References	16
8.2.	Informative References	16
	Authors' Addresses	17

1. Terminology

This document uses terminology discussed in [[RFC6388](#)] and [I-D.ietf-mpls-ldp-multi-topology]. Additional key terms and terminology are listed here:

- o PLR: The node upon which traffic is logically redirected onto the preset backup path is called Point of Local Repair (PLR).
- o MP: The node upon which the backup path and the primary path merge is called the Merge Point (MP).
- o N: The node which is protected by the backup path.
- o P_n: The nodes on the backup path that protect node N.
- o MT-ID: A 16 bit value used to represent the Multi-Topology ID.
- o Default MT Topology: A topology that is built using the MT-ID default value of 0.
- o MT Topology: A topology that is built using the corresponding MT-ID.
- o MRT: Maximally Redundant Trees. A pair of trees where the path from any node X to the root R along the first tree and the path from the same node X to the root along the second tree share the minimum number of nodes and the minimum number of links.

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

To meet user demands, operators and service providers continue to deploy multicast applications using Multicast LDP (mLDP) across MPLS networks. For real-time applications, such as stock trading, on-line games, and multimedia teleconferencing, traditional node protection mechanisms (such as IGP-mLDP convergence based mechanisms) fail to adhere to the protection switching time which is required to minimize the interruption of applications.

Instead of relying on the IGP-mLDP convergence for failure protection,

pre-computing, establishing a backup path before the failure and switching over to the backup path when the protected node fails is a better solution.

But two major challenges exist with this aforementioned solution. The first is how to compute an absolutely disjointed backup path for each node in a multicast tree; the second is how to signal and setup the backup path.

For a primary LDP P2MP LSP, there are several methods to choose for a backup path:

- o The use of an RSVP-TE P2P tunnel as a logical out-going interface, which consequently utilizes the mature high availability technologies of RSVP-TE.
- o The use of an LDP P2P LSP as a packet encapsulation, wherein complex configuration of P2P RSVP-TE can be skipped.
- o Creating a P2MP backup LSP according to IGP's loop-free alternative route.
- o Using multi topology technology, wherein the backup topology can be either statically configured or dynamically computed/ signaled using mechanisms specified in the draft of [I-D.ietf-rtgwg-mrt-frr-architecture].

When the backup path is present, there are two options for packet forwarding and protection switchover:

- o Option 1 The traffic sender transmits the stream on both the primary and backup path. Once the local traffic receiver detects a failure, the switchover will be relatively fast. However, the disadvantage of this method is that it consumes bandwidth because duplicate traffic will be sent on the primary and backup paths.
- o Option 2 The traffic sender transmits only on the primary path. Although bandwidth resource usage is minimized, cooperation is required to provide adequate switching times and to minimize high-layer application impact.

Ideally, if the switching time performance can be equal or better than that of Option 1, it is reasonable to choose option 2 to avoid bandwidth wastage. Some recommendations in this document are based upon this consideration.

Note that the computation and configuration of the primary topology and backup topology are out of the scope of this draft. The

algorithm can be LFA based, MRT based, or based off of any other algorithms/method available including the static and offline tools. In addition, detecting failure is also outside the scope of this document.

Compared to a P2P LSP based solution (specified in the draft of I-D.wijnands-mpls-mldp-node-protection), this P2MP LSP based solution not only uses mLDP mechanisms for both the primary path and backup paths, but also avoids unnecessary packet duplication.

The remainder of this document specifies the signalling procedures and protocol extensions for the P2MP LSP based mLDP node protection solution which was briefly introduced above.

3.1. Requirements

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

- o Computation of a disjointed (link and node) backup path within the multicast tree
- o Minimization of protection convergence time
- o Minimization of operation and maintenance cost
- o Optimization of bandwidth usage
- o More protect scenario coverage

3.2. Scope

The method to detect failure is outside the scope of this document.

The protections of leaf and root nodes are also outside the scope of this document.

4. mLDP Node Protection Example

By using IGP-FRR or Multi Topology Routing (including the MRT MT routing), LDP can build the backup mLDP LSP among PLR, Pn, and MPs (the downstream nodes of the protected node). In the case of a large number of downstream nodes, this mechanism can avoid unnecessary packet duplication between PLR and the merge points.

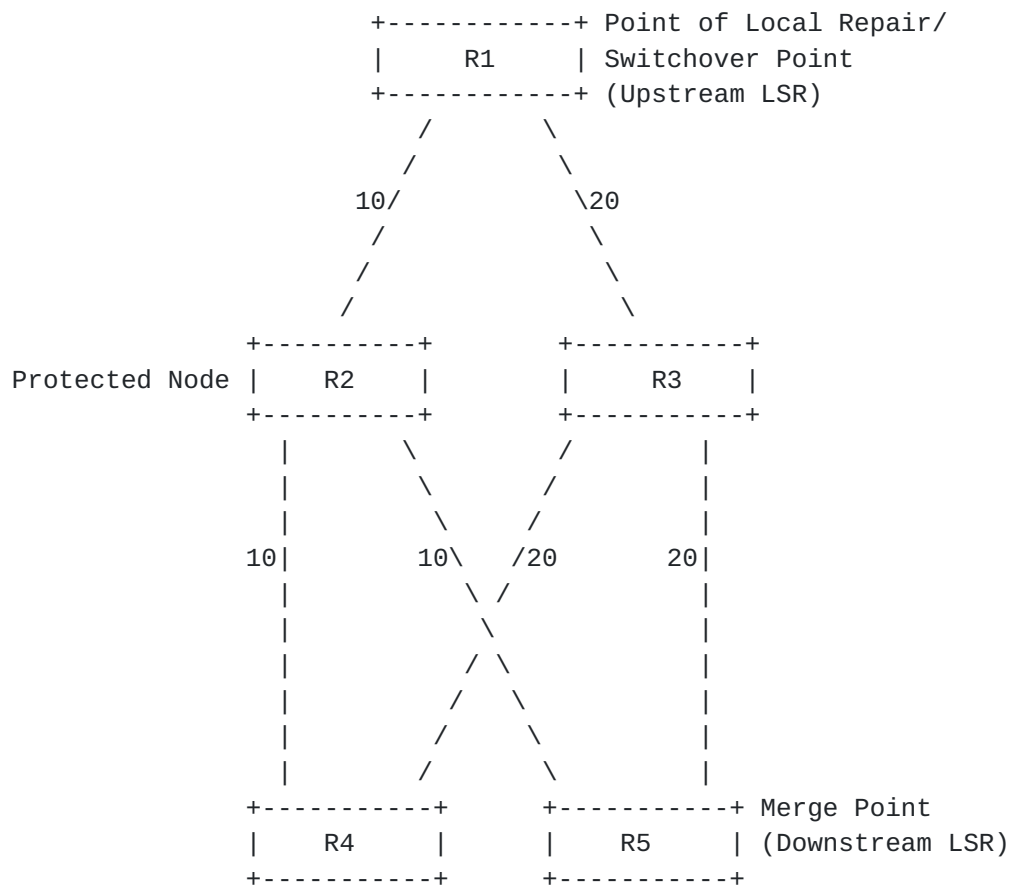


Figure 1: mLDP Local Protection using mLDP LSP Example

In Figure 1, R2 is on the preferential path from R4/5 to R1, 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. As an example, this figure takes R2 as the Protected Node though the Protected Node can be any Transit node of the mLDP LSP. (We assume that all the nodes in Figure 1 support this mLDP based node protection method, including Pn.)

The procedure of P2MP Based mLDP Node Protection is as follows:

- o As the Protected Node, R2 should announce its selected upstream node R1 to all its downstream nodes, which are R4 and R5 in this example. The node to protect can then be decided by local configuration or by its role(transit) in the mLDP LSP.
- o R4 and R5 can consider R1 as the root node of the backup mLDP LSP and can trigger the backup LSP signaling. In parallel, R4/R5 will bind the primary NHLFE(s) to both the backup and primary ILM entries, so that the traffic from the backup mLDP LSP can be

merged locally to the primary LSP.

- o PLR can distinguish primary LSP and backup LSP by the signaling procedure and can feed traffic on the primary path before failure. When R2 node fails, R1 quickly switches the traffic to the preset backup path.

In this scenario, if R2 is protected by two P2P LSPs as R1--R3--R4 and R1--R3--R5 (similar to the method in the draft of I-D.wijnands-[mpls-mlbp-node-protection](#)), the traffic will be duplicated on R1, and R3 will receive two streams. But, if R2 is protected by a mLDP LSP instead, R3 will only receive one stream, and thus there will be no packet duplication on R3.

5. Signaling Procedures for P2MP Based Node Protection

This section introduces the signaling procedures of P2MP LSP's node protection using P2MP-based backup LSP.

5.1. The Computation of the Backup Path

Obviously, the backup path can not go through the protected node N. This section discusses how to choose the backup upstream LSR to avoid N.

Firstly, find the candidate upstream LSRs as below:

- o MPs should preferentially choose the upstream LSRs on the shortest path as candidates, except node N. If no other upstream LSRs are on the shortest path, MPs should choose the next-hop on N's detour path as a candidate. The detour path can be an IGP-FRR path or other topology-based disjoint paths. The IGP-FRR path can be provided by LFA, U-Turn, etc. The disjoint path can be provided by MT, MRT (see details in the next section), etc. Choosing the candidates is a local decision and can be determined by configuration.
- o The Pn node MUST be chosen from the IGP next-hops on the shortest path toward PLR within the topology specified in the FRR mLDP FEC element by MT-ID field. The candidate upstream LSRs MUST not be the node N.

Thus, each node can choose one upstream node from the candidate upstream LSRs as its backup upstream LSR via the algorithm described in [\[RFC6388\] section 2.4.1.1](#).

5.2. The Procedures for Merge Point, Protected Node, Pn Node and PLR Node

[Editors Note - The procedures for P2MP Based Node Protection described in this document assumes the PLR is capable of node failure detection.]

We assume all the involved nodes have advertised their corresponding protection capabilities. And the following section specifies the signaling procedures of P2MP Based Node Protection.

5.2.1. Merge Point's Procedures

Each non-Ingress LSR determines its own upstream LSR and sends out a label mapping message, in accordance with the procedures documented in [[RFC6388](#)] without any additional extension. And its upstream LSR will propagate a new label mapping message to its upstream LSR. In such cases, the non-Ingress LSR is the MP node (as R4, R5 in Figure 1), MP's upstream LSR is the protected node (as R2 in Figure 1), and the protected node's upstream node is PLR (as R1 in Figure 1).

When one MP (as R4/R5 nodes in figure 1) receives the Notification, it individually determines its secondary path toward the PLR according to the IGP routes. The algorithms for choosing/computing the backup path can be LFA, MRT or others. After the backup upstream LSR is chosen, MP will send out a FRR Label mapping message, which includes the mLDP backup tree's key <PLR, protected-node, original-mLDP-FEC> and the MT-ID if the backup path is not in the default topology. Note that the label assigned for the primary path and the secondary path MUST be different to avoid having the MP feed primary traffic to its secondary path's downstream LSRs. In addition, the original-mLDP-FEC of the backup tree key is encoded in a special opaque value as introduced in [section 4.2.3](#).

5.2.2. Protected Node's Procedures

After the Protected Node (as R2 in Figure 1) determines its upstream LSR (as R1 in figure 1), it will send the information (PLR's indentify, mLDP FEC) via Notification messages to all its downstream nodes(MPs) immediately. If other LSRs become its downstream nodes later, it will send such announcements to its new MP(s).

5.2.3. Pn Node's Procedures

When one node receives such aforementioned FRR label mapping messages, if it is not the PLR, it can consider itself a Pn node and will choose its backup upstream node toward PLR on the corresponding topology's shortest IGP path. To avoid the backup LSP going through

the Protected Node, additional path selection rule(s) should be applied. A simple method is for the transit nodes to not choose the specified Protected Node as its upstream LSR for the backup LSP. Other methods, such as the not-via policy, are under study and will be added in the future. To make the primary and backup topologies rooted from PLR satisfy the 'maximum disjointed' requirement, they can either be configured through static configurations or be signaled dynamically through other mechanisms such as MRT.

5.2.4. PLR Node's Procedures

When PLR(R1) receives the FRR label mapping message, it can identify that it is the PLR by the mLDP backup FEC elements. Thus, it will decode the special opaque value (which contains the primary mLDP FEC element, introduced in [section 4.2.3](#)) and generate the backup forwarding entry for the specific LSP, which is identified by the root address and opaque value in the special opaque value. It will also bind the backup forwarding state to the specific primary entry, which is indicated by the Protected Node address in the message. Note that there might be more than one backup forwarding entry for a specific protected node.

When failure is detected by PLR, it will switch the traffic to the backup paths. MP will also locally choose which traffic to receive and merge this traffic back to the primary LSP. The switchover manner on PLR is specified in detail in the later section of this document.

5.3. PLR Switching Over Considerations

This section provides two methods for Switchover when failure occurs:

- o Option 1: If PLR cannot differentiate link and node failure, MP must take the responsibility to drop one of the two duplicate traffics when failure is detected. In this case, the Node Failure Required Flag (in the P2MP Based MP Node Protection Status Element) must be set as 'N'. PLR will switch the traffic to the backup path when failure is detected, and MP will drop traffic on the backup path until it sees N fail.
- o Option 2: If PLR can differentiate link and node failure, PLR MUST NOT switch the traffic to the backup path until it detects the node N's failure. In this case, the Node Failure Required Flag, in the P2MP Based MP Node Protection Status Element, must be set as 'Y'.

5.4. The Procedures after the Reconvergence

When Merge Point(s) see the next hop to Root changed, it/they will advertise the new mapping message(s), and the traffic will re-converge to the new primary path. MP will then withdraw the backup label after the re-convergence. Pn will delete the specified backup LSP just as in the procedure of deleting normal P2MP LSP. And the entire backup P2MP LSP will be deleted when the node MP leaves the backup P2MP LSP.

5.5. Considerations for MP2MP LSP Node Protection

When a MP2MP LSP node needs to be protected, it can be treated with the same p2mp LSP node protection procedures for each forwarding direction.

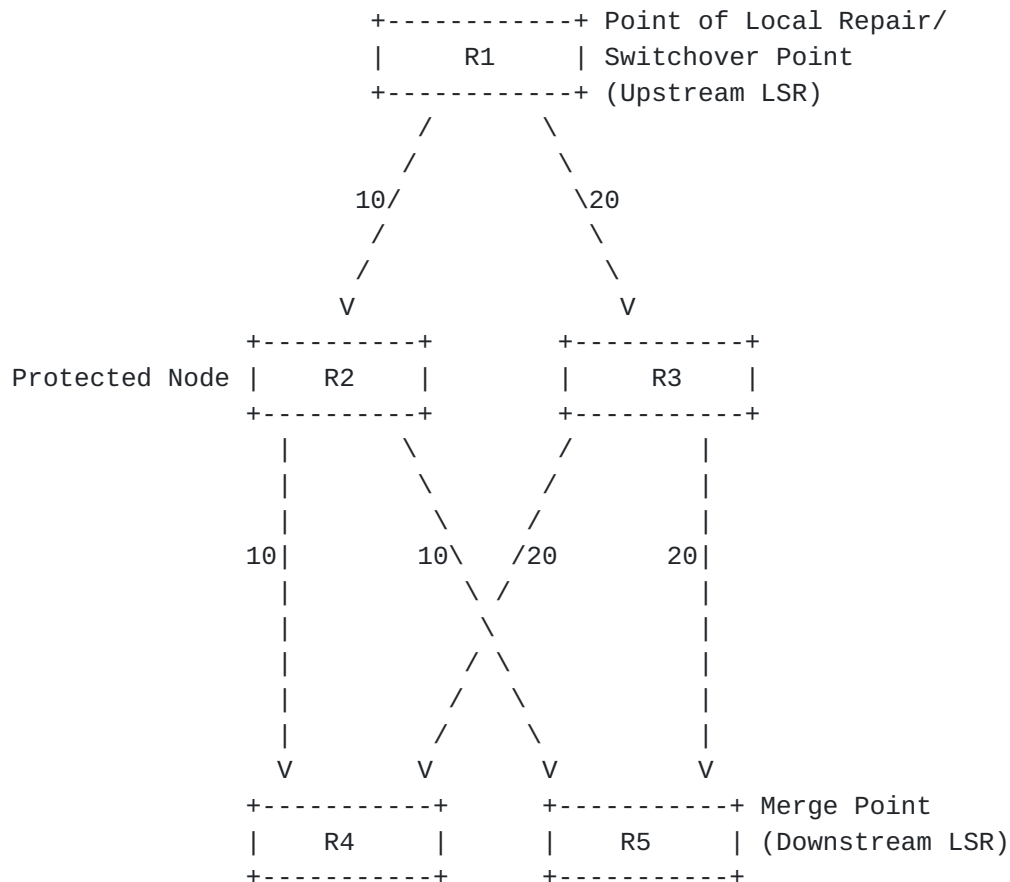


Figure 2: MP2MP Example (R1 is the PLR)

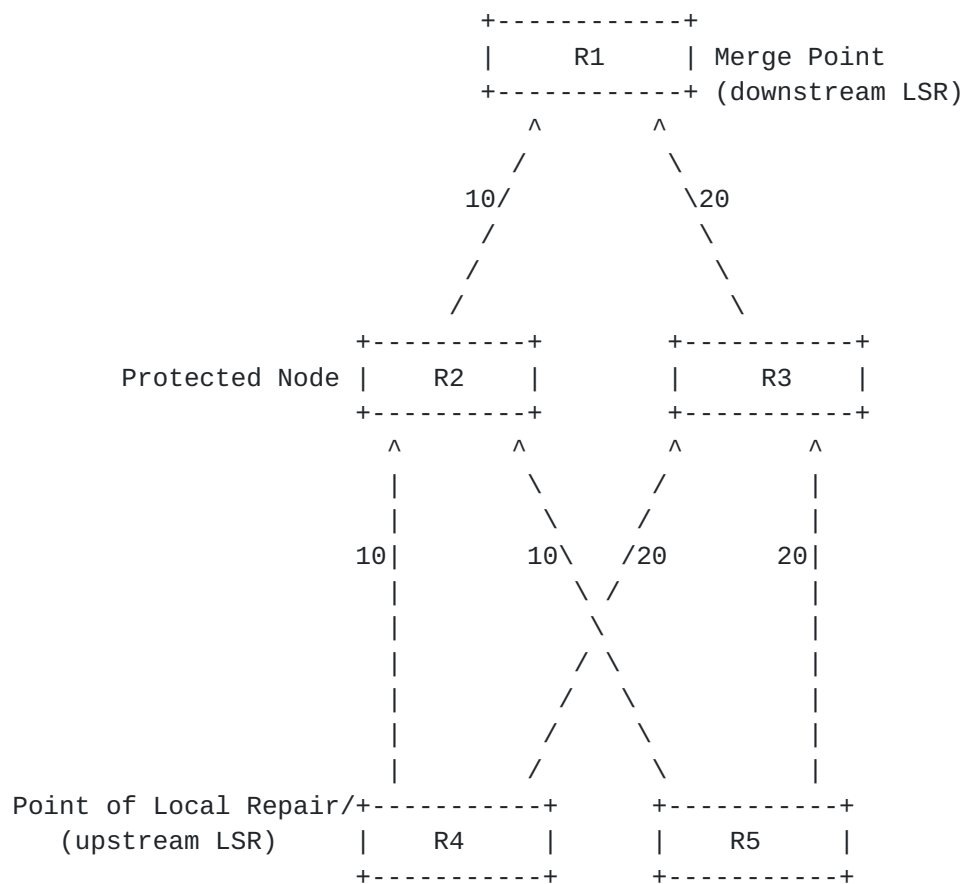


Figure 3: MP2MP Example (R4 is the PLR)

For each direction flow, the MP, PLR and MP nodes use the P2MP node protection procedures with the following additional considerations:

5.5.1. MP Node Procedure

MP sends a backup label mapping message containing MP2MP downstream FRR FEC elements. When PLR receives a backup label mapping message with a MP2MP downstream flag, it sends the backup label mapping message with mp2mp upstream FRR FEC elements to Pn and then finally to MPs. This procedure just follows the normal MP2MP LSP procedure.

For the forwarding entries, MP node binds its primary MP2MP downstream NHLFE entry to backup MP2MP downstream ILM entry and binds its backup MP2MP upstream NHLFE entry to primary MP2MP upstream ILM entry.

For the forwarding entries, MP node binds its primary MP2MP downstream NHLFE entry to backup MP2MP downstream ILM entry and binds its backup MP2MP upstream NHLFE entry to primary MP2MP upstream ILM

entry.

5.5.2. PLR Node Procedure

PLR node binds its backup MP2MP downstream NHLFE entry to primary MP2MP downstream ILM entry and binds its primary MP2MP upstream NHLFE entry to backup MP2MP upstream ILM entry.

5.5.3. Switchover Procedure

When the protected node fails, both the affected downstream and upstream nodes function as PLR and switch the downstream flow and upstream flow to their respective backup paths.

5.6. Protocol Extensions for mLDP Based Node Protection

5.6.1. mLDP Based MP Protection Capability Parameter TLV

A new Capability Parameter TLV is defined as mLDP Based MP Protection Capability for node protection. The following is the format of this new Capability Parameter TLV:

```

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
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|1|0| mLDP Based MP Prot.(IANA) |           Length (= 2)           |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|S| Reserved                    |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

S: As specified in [[RFC5561](#)]

Figure 4: mLDP Based MP Protection Capability

This is an unidirectional capability announcement.

An LSR, which supports the mLDP based protection procedures, should advertise this mLDP Based MP Protection Capability TLV to its LDP speakers. Without receiving this capability announcement, an LSR MUST NOT send any message including the mLDP Based MP Node Protection Status Element and mLDP Backup FEC Element to its peer.

Capability Data might be needed to distinguish the capabilities of different nodes, such as PLR, MP, N, Pn and so on.

5.6.2. mLDP Based MP Node Protection Status Elements

A new type of LDP MP Status Value Element is introduced for notifying upstream LSR information. It is encoded as follows:

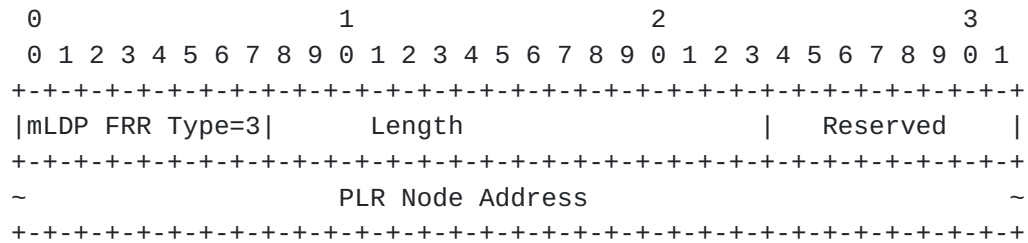


Figure 5: FRR LDP MP Status Value Element

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

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

PLR Node Address: The host address of the PLR Node.

5.6.3. mLDP Backup FEC Element Encoding

A new type of mLDP backup FEC Element is introduced for notifying upstream LSR information. It is encoded as follows:

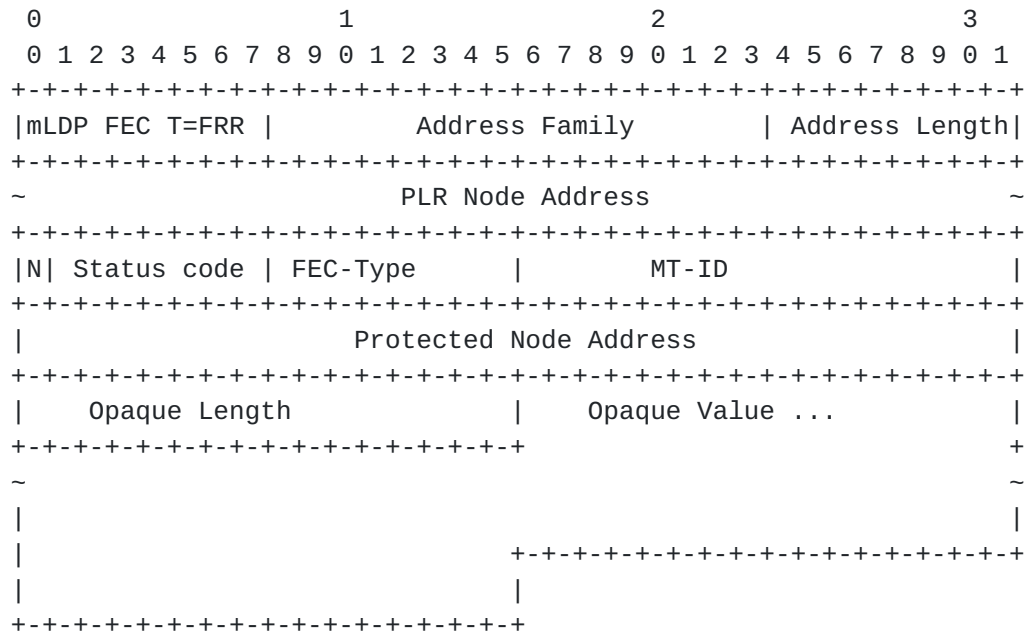


Figure 6: mLDP Backup FEC Element

mLDP FEC Type-FRR: Type 5 (to be assigned by IANA)

Length: If the Address Family is IPv4, the Address Length MUST be 9;
if the Address Family is IPv6, the Address Length MUST be 33.

Status code: 1 = Primary path for traffic forwarding
2 = Secondary path for traffic forwarding

FEC-Type: 6 = P2MP FEC type
7 = MP2MP-up FEC type
8 = MP2MP-down FEC type

PLR Node Address: The host address of the PLR Node.

Protected Node Address: The host address of the Protected Node.

N Bit: Node Failure Required Flag, the occasion of switching traffic's on PLR

1 = 'Y', switch traffic to backup path only when PLR detects the node failure

0 = 'N', switch traffic to backup path when PLR detects failure

Opaque Length: The length of the opaque value, in octets.

Opaque Value: One or more MP opaque value elements, which is the same definition in [RFC6388].

For the FRR mLDP FEC element, the Opaque Value MUST be encoded as the Recursive Opaque Value, which is defined in [RFC6512].

The value

primary

fields of the Recursive Opaque Value contain the original

path's mLDP FEC element.

The encoding for this Recursive Opaque Value, as defined in [RFC6512], is shown in Figure 5.

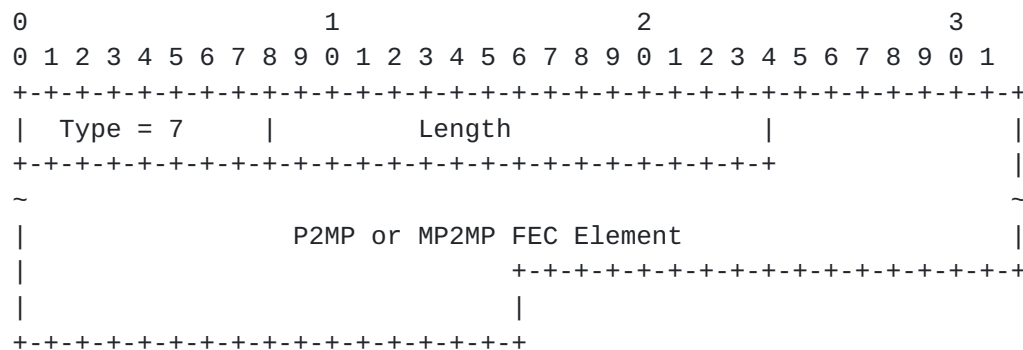


Figure 7: Recursive Opaque Value, defined in [RFC6512]

The Opaque Value is encoded by MP node and decoded by PLR. Other nodes MUST NOT interpret the opaque value at all.

5.7. IANA Considerations

This memo includes the following requests to IANA:

- o mLDP Based MP Protection Capability
- o mLDP FRR types for LDP MP Status Value Element
- o mLDP FEC FRR Element type

6. 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.

7. Acknowledgements

We would like to thank Nicolai Leymann and Daniel King for their valuable suggestions to this draft. We also would like to thank Robin Li, Lujun Wan, Kenji fujihiara, Martin Vigoureux, Yaacov Weingarten and Eric Osborne for their comments and suggestions to the draft.

8. References

8.1. 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.
- [RFC6388] Wijnands, IJ., Minei, I., Kompella, K., and B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", [RFC 6388](#), November 2011.
- [RFC6512] Wijnands, IJ., Rosen, E., Napierala, M., and N. Leymann, "Using Multipoint LDP When the Backbone Has No Route to the Root", [RFC 6512](#), February 2012.

8.2. Informative References

- [I-D.ietf-mpls-ldp-multi-topology]
Zhao, Q., Fang, L., Zhou, C., Li, L., and K. Raza, "LDP Extensions for Multi Topology Routing", [draft-ietf-mpls-ldp-multi-topology-08](#) (work in progress), May 2013.

[I-D.wijnands-mpis-mldp-node-protection]

Wijnands, I., Rosen, E., Raza, K., Tantsura, J., Atlas, A., and Q. Zhao, "mLDP Node Protection", [draft-wijnands-mpis-mldp-node-protection-04](#) (work in progress), June 2013.

[I-D.ietf-rtgwg-mrt-frr-architecture]

Atlas, A., Kebler, R., Envedi, G., Csaszar, A., Tantsura, J., Konstantynowicz, M., White, R., and M. Shand, "An Architecture for IP/LDP Fast-Reroute Using Maximally Redundant Trees", [draft-ietf-rtgwg-mrt-frr-architecture-02](#) (work in progress), February 2013.

[I-D.enyedi-rtgwg-mrt-frr-algorithm]

Atlas, A., Envedi, G., Csaszar, A., and A. Gopalan, "Algorithms for computing Maximally Redundant Trees for IP/LDP Fast- Reroute", [draft-enyedi-rtgwg-mrt-frr-algorithm-02](#) (work in progress), October 2012.

Authors' Addresses

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

Email: quintin.zhao@huawei.com

Tao Chou
Huawei Technology
156 Beiqing Rd
Haidian District, Beijing 100095
China

Email: tao.chou@huawei.com

Boris Zhang
Telus Communications
200 Consilium Pl Floor 15
Toronto, ON M1H 3J3
Canada

Phone:
Email: Boris.Zhang@telus.com

Emily Chen
2717 Seville Blvd, Apt 1205
Clearwater, FL 33764
US

Email: emily.chen220@gmail.com

