Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 14, 2012

Q. Zhao
E. Chen
T. Chou
Huawei Technology
D. King
Old Dog Consulting

March 13, 2012

Protection Mechanisms for Label Distribution Protocol P2MP/MP2MP Label Switched Paths
draft-zhao-mpls-mldp-protections-02.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 minimize switching times in the event of failure.

This document outlines the requirements, procedures and extensions to facilitate mLDP Point-to-Multipoint (P2MP) and Multipoint-to-Multipoint (MP-to-MP) LSP protection within an MPLS network.

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 September 14, 2012.

Table of Contents

1. Terminology

This document uses terminology discussed in [RFC5036] and [MT-LDP]. Additionally the following section provides further explanation for key terms and terminology:

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

- N: The node being protected.

- Pn: The node(s) on the backup path for protecting node N.

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

- FD: The node that detects the failure on primary path, and then triggers the neccessary action for traffic protection is called Failure Detector (FD). Either traffic sender or receiver can be the FD, depending on which protection mode has been deployed. Further specification is provided in later sections.

- SP: The node where the traffic is physically switched/duplicated onto the backup path is called Switchover Point (SP). In multicast scenarios, PLR and SP can be two different nodes. Further specification is provided in later sections.

- T-LDP: Targeted LDP session.

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 Multicast LDP (mLDP) across MPLS networks. In certain key scenarios, conventional 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 subsequently

higher-layer applications, include the pre-computation and establishment of a backup path. Once a failure has been detected on the primary path, the traffic will be transmitted across the back-up path.

However, two major challenges exist with the existing 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, resource consumption and network efficiency.

For a primary LDP P2MP/MP2MP LSP, there are several methods to set up a backup path, these include:

- The use of an RSVP-TE P2P tunnel as a logical out-going interface, consequently utilize the mature high availability technologies of RSVP-TE.

- The use of an LDP P2P LSP as a packet encapsulation, so that the complex configuration of P2P RSVP-TE can be skipped.

- Creating a P2MP/MP2MP backup LSP according to IGP's loop-free alternative route. Comparing to using P2P LSPs, this solution can

prevent unnecessary packet duplication on common links.

- Creation of Multiple Topology (MT) LSP using an entirely disjointed topology.

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

- 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 as duplicate traffic will be sent on the protection and backup path.

- 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 minimise high-layer application impact.

Ideally if switching time performance for Option 2 can be closer to the Option 1, it is reasonable to choose it to avoid bandwidth wastage. The recommendations of this document are based on this point of view.

This document provides several ways to setup the backup path for mLDP LSP, including P2P based mLDP node protection, P2MP based mLDP node protection, and MT based end-to-end protection. The goal is to build a reliable umbrella to against traffic black hole.

Note that the backup path computation is out of the scope of this draft, the algorithm can be either LFA or any other algorithms available including the offline tools. Besides, how to detect failure is also outside the scope of this document, the mechanism can be bidirectional or unidirectional forwarding detection for link or target object.

3.1. Requirements

A number of requirements have been identified that allow the optimal

set of mechanisms to developed. These currently include:

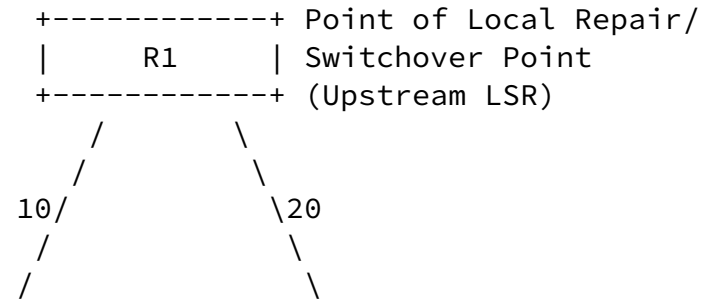
- Computation of a disjointed (link and node) backup path within the multicast tree;
- Minimization of protection convergence time;
- Minimization of operation and maintenance cost;
- Optimization of bandwidth usage;
- Minimization the impact on the existing network deployment.

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. mLDP Node Protection using P2P LSPs

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 recovering the failure on the destination node of the P2P backup LSP. Thus, this section provides an extra method to protect node using an P2P LSP.



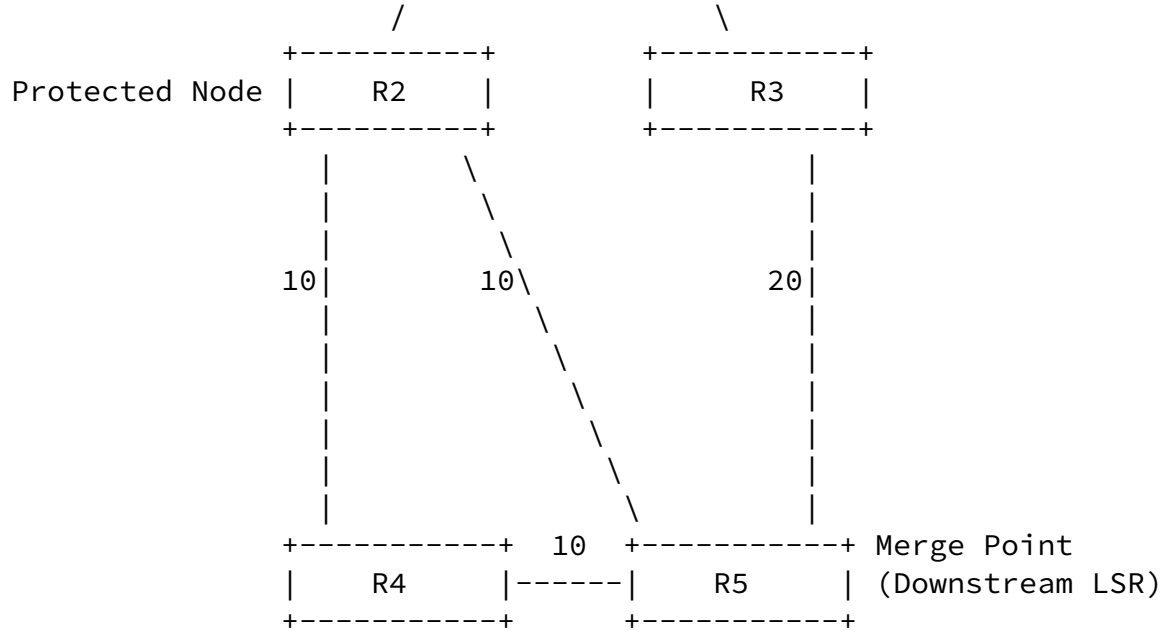


Figure 1: P2P Based mLDP Node Protection Example

In Figure 1 (P2P Based mLDP Node 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 <X,Y>, will be established according to the IGP preferential path as R1-- R2--R4/R5. R4's backup P2P path is R1--R3--R5--R4. R5's backup P2P path is R1--R3--R5. (We assume that all the nodes in Figure 1 support the P2P Based mLDP Node Protection capability.)

The procedure for P2P Based mLDP Node Protection is as follows:

The MP(s) (in this example R4 or R5) sends label mapping message with protection information to the the protected node N(in this example R2). Node N notifies all the dowstream information to the node PLR(in this example R1). PLR seletes P2P LSP toward each MP(s) as backup path encapsulation. When PLR detects N failed, PLR switches traffic to MP(s) over corresponding backup path(s). PLR will stop the backup traffic forwarding after relative MP(s) finish convergence procedure.

There are two methods for PLR to switch over the traffic:

o Option 1:

If PLR can differentiate the link and node failure, such as binding two BFD sessions on link and node, then it can feed the traffic to the node protection path only when the node failure is detected. Only N is required to maintain BFD session with the PLR, it does not need to maintain the large number of BFD sessions with MPs.

- Option 2:
  If the PLR can not differentiate between link and node failure, it should feed the traffic to both link protection path and node protection path at the same time, and MP must take the responsibility to drop one stream. In this case, N must maintain BFD sessions with PLR and all the MPs.

[Editors Note - The authors recommend the first method to save BFD resource usage, the details are specified in section 4.1.2. This Editors note and remaining options will be adjusted once we get more feedbacks from users.]

Additionally there are two methods for the backup label cleanup between the MP and PLR:

- Option 1:
  The PLR can delete the backup label by receiving a withdraw notification message from MP through a T-LDP between PLR and MP. This method needs PLR maintains a huge number of T-LDP sessions with MPs.

- Option 2:
  The second method requires cooperation between PLR and MP by a synchronous timer. The MP will synchronize the timers to the PLR and hold the backup label resource until its local timer expires. In addition the PLR will delete the backup label when its local timer expires.

[Editors Note - The authors recommend the second method in order to save T-LDP resource usage, the details are specified in section 4.1.3. This Editors note and remaining options will be adjusted once we get more feedbacks from users.]

4.1. Signaling Procedures for P2P Based Node Protection

4.1.1. P2P Based Node Protection Procedure Example

[Editors Note - This section introduces the procedures for P2P Based Node Protection based on the two options recommended above.]

The following in this section demonstrates the signaling and procedures for P2P Based Node Protection. Note that STEP5 and STEP6 should be acted at the same time.

STEP1 MP's procedures for setting up backup path:
Take R4 for example, which acts as a node MP. R4 determines N (in this example R2) as its upstream LSR, and then sends label mapping message to R2 including label L4 for <X,Y>. This label mapping message MUST include a new LDP MP Status Value Element, which includes a backup label and a reserve-time. The backup label L4' is assigned by MP for PLR's backup path. In order to avoid upgrading the MP's hardware to support filtrating traffic, this backup label can be same with L4. This reserve-time is the vlaue of the reserve-timer on MP, which can be configured, and it is recommend to be longer than the IGP convergence time and LDP MBB procedure time.

STEP2 N's procedures for setting up backup path:
When R2 receives the mapping message from MP, it determines R1 as its upstream LSR and send a label mapping message with label L2 to R1. Besides, R2 transfers all the new LDP MP Status Value Elements, received from its downstreams, to R1 by a notification message.

STEP3 PLR's procedures for setting up backup path:
When R1 receives this notification message and label mapping message from R2, R1 creates two forwarding paths: the primary forwarding state is <X,Y,L2> , and the secondary forwarding states is <X,Y,L4'>. Note that there might be more than one secondary forwarding states, such as <X,Y,L5'>.

STEP4 PLR's procedures when node N fails:
Once the node PLR(R1) detects the node N(R2) failure, before protocol converging, R1 will switch the traffic to MP(R4) over P2P LSP, with inner label as L4', and outer label as the P2P backup LSP R1--R3-- R5--R4. Note that the PLR MUST NOT switch traffic to backup path until it detects N failure. R1 will also duplicate the traffic to other MPs(for example R5) over relative P2P LSPs. Meanwhile, the node PLR will create a reserve-timer with the reserve-time value in the notification message for each N's MP respectively. Note that once the primary forwarding state is removed, the secondary forwarding state MUST be deleted too.

STEP5 MP's procedures after convergence:
MP will also create the reserve-timer when it detects the

network convergence. MP(s) will remove the previous forwarding state after a new path to root is created or MBB

procedure finishes. Note that MP MUST hold the old bakcup label resource (L4') until its local reserve-timer expires.

STEP6 PLR's procedures after convergence:
When R1's local reserve-timers expire, it will stop the traffic on the backup paths and remove the secondary forwarding states.

Note that the mLDP Local Protection mechanism can be used in any part of the mLDP LSP other than the ingress and egress nodes. In other words, R1 can be either Ingress or Transit node, R4/R5 can be either Transit or Egress node.

4.1.2. PLR Switching Over Considerations

PLR switching over method depends on its failure detection mode. If PLR switches traffic when a link failure happens, MP may receive reduplicate traffic because the node N still feeds the traffic to MP. This is a problem if MP accepts these reduplicate traffic. There are two optional methods to solve this problem.

4.1.2.1. Single Feed Mode When Node Failure Detection is Supported

In this method, the node PLR MUST be capable of differentiating link and node failure, and MP will not receive reduplicate traffic. It does not need MP's cooperation in such case, and the new label assigned to PLR has no need to be different with the one assigned to N.

MP also needs to send label mapping message with a new LDP MP Status Value Element to the node N. Especially, the MP Status TLV's Node Failure Required Flag need to be set as 'Y'. The node N transfers this MP Status TLV to the node PLR by notification message.

In such case, the node PLR MUST NOT switch the traffic to the backup path until it detects the node N failure.

4.1.2.2. Duel Feed Mode When Node Failure Detection is Not Supported

In this method, MP must support traffic filtering. Thus, PLR doesn't need to differentiate link and node failure becuase it can cooperate with the node MP. The node MP may receive reduplicate traffic and MUST drop the backup traffic when the node N doesn't fail. To support this manner, MP MUST assign PLR a new label Lp, which is different from the label Ln assigned for N.

The node MP send label mapping message with a new LDP MP Status Value Element to the node N. The label Ln is encoded in the label TLV and

the Backup Label Lp is encoded in the MP Status TLV. Besides, the MP Status TLV's Node Failure Required Flag need be set as 'N'. The node N transfers this MP Status TLV to the node PLR by notification message.

If the MP Status TLV's Node Failure Required Flag is 'N', before failure occurres, the label Ln is activated and Lp is deactived on the node MP. The PLR will switch traffic to the backup path when it detects the failure, no matter whether it is link failure or node failure.

The node of MP needs to detect whether the node N is failure or not. This is why N needs to maintain BFD sessions with all the MP nodes. If N fails, MP should deactivate the primary forwarding state of label Ln, and active the secondary forwarding state of label Lp. So the MP will receive the traffic on the backup path, and drop the traffic on the primary path. Otherwise, if it is link failure, MP will keep the backup label Lp deactivate and the traffic PLR switching to the backup path is dropped by MP.

4.1.3. Backup Path Cleanup Considerations

In order to prevent traffic duplication and unnecessary traffic lost, PLR and MP should delete the label at the same thime. There are two methods to ensure this.

4.1.3.1. Timer Synchronization Mode

In this mode, a reserve-timer is needed on both MP and PLR. The value of this timer can be configured on MP, and it is recommend to be longer than the IGP convergence time and mLDP MBB procedure time.

Once the reserve time is configured, the new LDP MP Status Value Element in label mapping message MUST include this time value and the Delete Flag MUST be set as 'implicit-delete', mentioned in Section 4.4.

The Node N transfers this time value and flag to the node PLR by the notification message. When the node PLR receives this notification message, PLR checks the Delete Flag and uses this time value as PLR's reserve timer if Delete Flag equals 'implicit-delete'.

PLR will create the reserve-timer when it detects the failure and switch traffic to the backup path. PLR will remove the secondary forwarding state when this reserve-timer expire.

MP will create the reserve-timer when it detects the network convergence. MP will remove the old forwarding state after a new

path to root is created or MBB procedure finishes. Noted that, MP MUST hold the old label resource not return it to the free pool until its reserve-timer timeout.

4.1.3.2. T-LDP Mode

In this mode, an automatic setting up T-LDP is needed between MP and PLR. It requires the node MP and PLR MUST be capable of setting up T-LDP as documented in [RFC5036].

If MP need a T-LDP, the new LDP MP Status Value Element TLV in label mapping message MUST set its Delete Flag as 'explicit-delete'.

The Node N transfers this flag to the node PLR by the notification message. When the node PLR receives this notification message, PLR checks the Delete Flag and trigger a T-LDP toward to MP if flag equals 'explicit-delete'.

MP will remove the old forwarding state after a new path to root is created or MBB procedure is end. Meanwhile, MP MUST send a notification message to PLR through T-LDP. This message include a new LDP MP Status Value Element TLV, whose status code is 'withdraw' defined.

PLR will not remove the secondary forwarding state until it receive

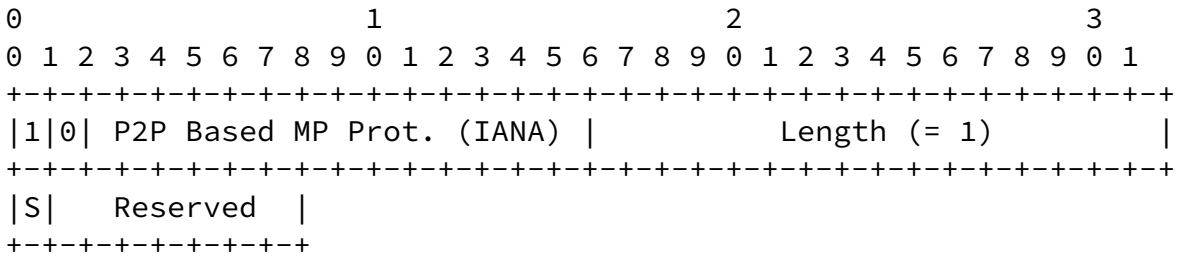
the notification message with 'withdraw' status code.

Noted that if the primary forwarding state is removed, the secondary forwarding state need be deleted no matter which delete method is used.

4.2. Protocol Extensions for P2P Based Node Protection

4.2.1. P2P Based MP Protection Capability Parameter TLV

A new Capability Parameter TLV is defined as P2P Based MP Protection Capability. Following is the format of this new Capability Parameter TLV:



S: As specified in [RFC5561]

This is an unidirectional capability announced.

An LSR, which supports the P2P based protection procedures, should advertise this P2P Based MP Protection Capability TLV to its LDP speakers.

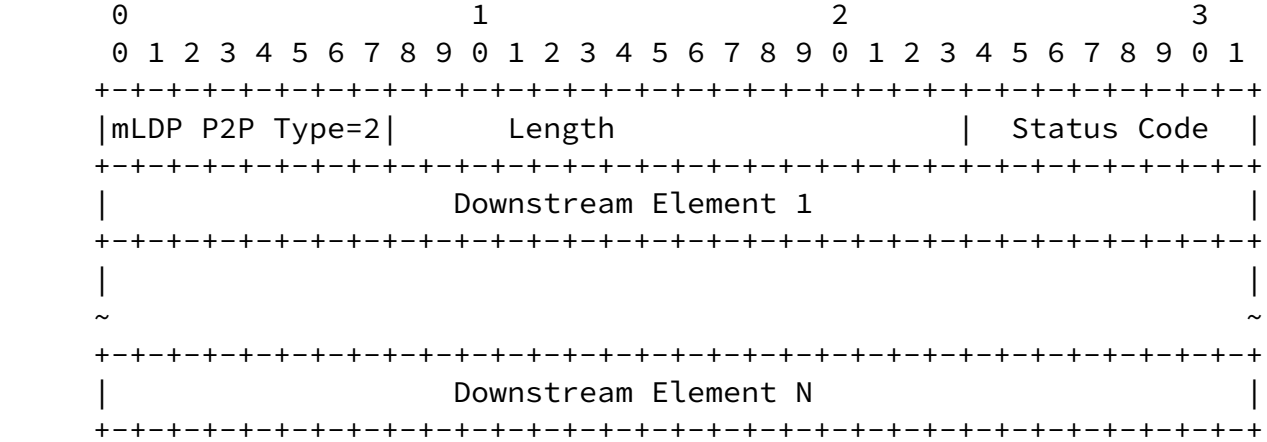
An LSR can consider that has no P2P Based MP Protection Capability if it dose not announce its capability. An LSR MUST NOT send any message including the new LDP MP Status Value Element TLV to its peer, which dose not have the P2P Based MP Protection Capability.

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

4.2.2. P2P Based MP Node Protection Status Element

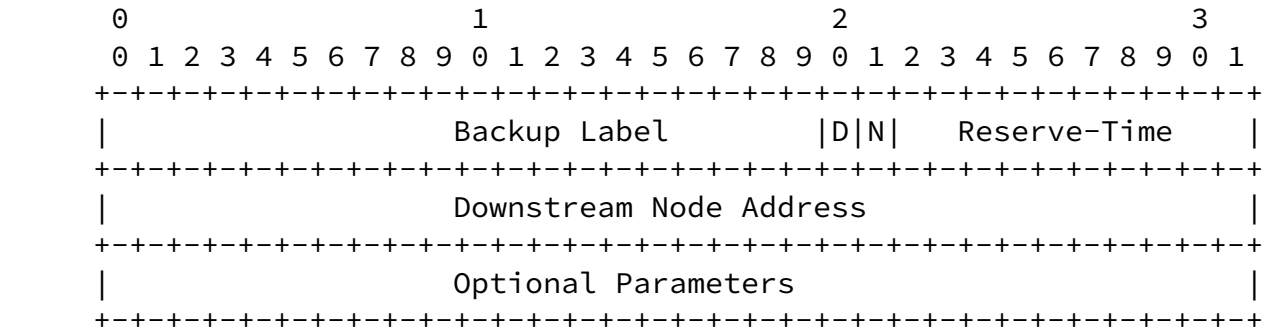
A new type of LDP MP Status Value Element is introduced, for notifying downstream LSR information, including respective labels and

other parameters. It is encoded as follows:



Status Code: 1 = Advertise the existing downstream LSRs
2 = Withdraw the deleted downstream LSRs

The Downstream Element is encoded as follows:



Backup Label: The label assigned by MP for PLR

D Bit: Delete Flag, The type of deleting backup label:
    1 = 'explicit-delete', delete by MP's notification message
        through T-LDP
    0 = 'implicit-delete', delete by reserve-timer expire
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
Downstream Node Address: Downstream node's LSR-ID address
Res-time: The time of MP's reserve-timer, synchronizing to PLR.
          It is effective when D bit set as 'implicit-delete'
          and MUST be ignored when D bit set as 'explicit-delete'.

5. mLDP Node Protection using P2MP LSPs

By using IGP-FRR, LDP can build the backup mLDP LSP among PLR, the protected node, and MPs (the downstream nodes of the protected node). In the cases where the amount of downstream nodes are huge, this mechanism can avoid unnecessary packet duplication on PLR, so that protect the network from traffic congestion risk.

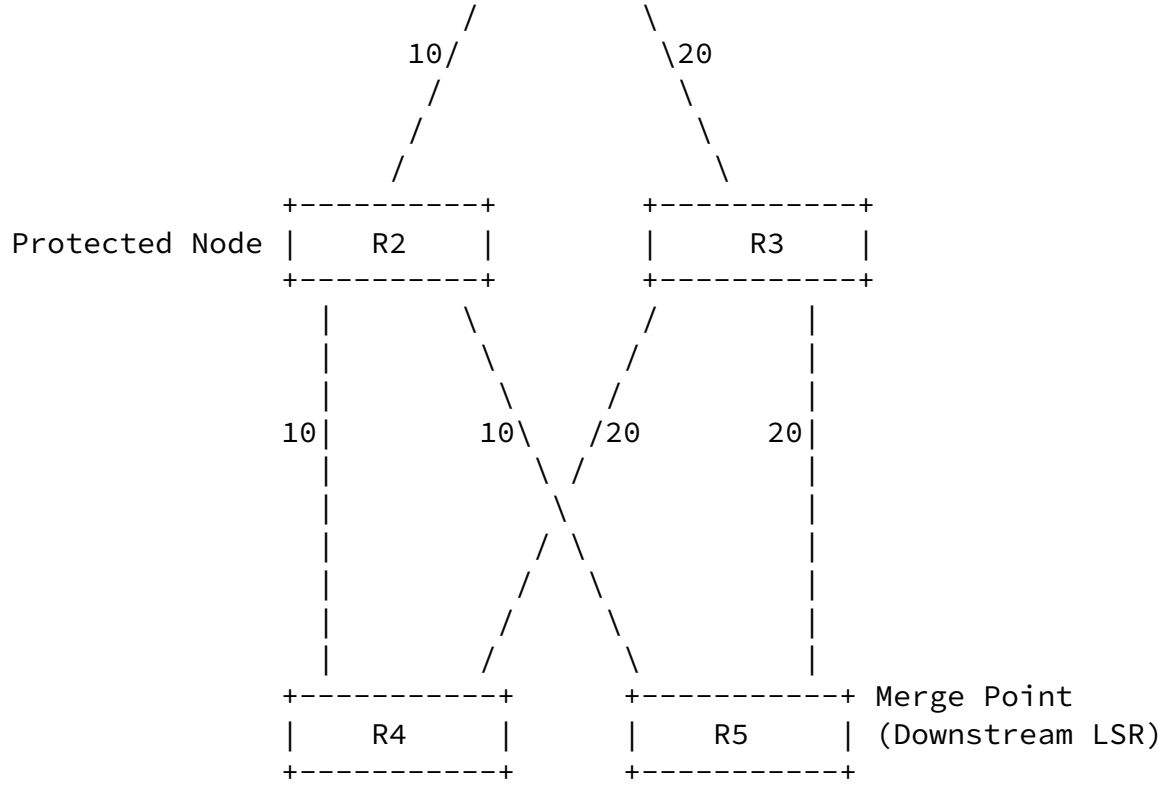


Figure 2: mLDP Local Protection using P2MP LSP Example

In Figure 2 (P2MP Based mLDP Local Protection Example), 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. This section will take the Protected Node as R2 for example, actually the Protected Node can be any Transit node of the mLDP LSP. (We assume that all the nodes in Figure 2 surpport this P2MP based node protection method, including Pn.)

The procedure for P2P Based mLDP Node Protection is as follows:

As the Protected Node, R2 will announce its selected upstream node R1 to all its downstream nodes, which are R4 and R5 in this example, when it receives thses label mapping messages.

R4 and R5 can consider R1 as the root node of the backup mLDP LSP, and trigger the backup LSP signaling. In parallel, R4/R5 will bind the primary NHLFE(s) to both the backup and primary ILM entry, so that the traffic receiving from backup mLDP LSP can be

merged locally to the primary LSP.

The primary LSP and backup LSP are differentiated by the signaling procedure, so normally PLR can only feed traffic on the primary path. When R2 node fails, R1 will switch the traffic to the preset backup path quickly.

In this scenario, if R2 is protected by two P2P LSPs as R1--R3--R4 and R1--R3--R5, the traffic will be duplicated on R1, and R3 will receive two streams. If R2 is protected by mLDP LSP instead, R3 will only receive one stream, and the packet duplication will be done on R3.

5.1. Signaling Procedures for P2MP Based Node Protection

5.1.1. P2MP Based Node Protection Procedure Example

[Editors Note - This section introduces the procedures for P2MP Based Node Protection based on the PLR being capable for node detection.]

We assume all the involved nodes have advertised their corresponding capabilities. And the following in this section demonstrates the signaling and procedures for P2MP Based Node Protection.

STEP1 N's procedures for setting up backup path:
MP determins N as its upstream and sends label mapping for to N, following the procedures as documented in [RFC6388] without any extention. When the Protected Node (R2) receives these label mapping messages and determines its upstream LSR (R1), it will notify to all its downstream nodes immediately. If there are other LSR(s) becoming its downstream node(s) later, it will do the announcement for the new downstream node(s).

STEP2 MP's procedures for setting up backup path:
When the Merge Point (R4/R5) receive the notification, they individually determine the primary and secondary paths toward R1 according to the IGP-FRR results. Then they will send out label mapping messages including an LDP MP Status TLV that carries a FRR Status Code to indicate the primary path and secondary path. The backup path is uniquely identified by root address, opaque value, PLR Node address, and Protected Node address. Noted that, the label assigned for primary parh and secondary path MUST be different to avoid the MP feeding the primary traffic to its secondary path's downstream LSRs.

STEP3 Pn's procedures for setting up backup path:
When the transit nodes of the secondary LSP receive the FRR label mapping message, they can easily consider it as a new mLDP LSP establishment, and follow the existing protocol procedures. The modification for these nodes is dealing with the FRR FEC, which is identified by root address, opaque value, PLR address, and Protected Node address.
To avoid the backup LSP going through the Protected Node, additional path selection rule(s) can be applied. A simple method is that the transit nodes can not choose the specified Protected Node as its upstream LSR on the secondary LSP. Other methods, such as not-via policy, are under study, and will be added in the future.

STEP4 PLR's procedures for setting up backup path:
When the Point of Local Repair (R1) receives the FRR label mapping message, it will generate the backup forwarding entry for the specific LSP, which is identified by the root address and opaque value in the message, and 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 entries for a specific protected node.

STEP5 PLR's procedures when node N fails:
When failure is detected by PLR, it will switch the traffic to the secondary path. MP will also locally merge the traffic back to the primary LSP. The switchover manner on PLR is specified in the later section.

STEP6 Procedures after network re-converges:
When Merge Point(s) see the next hop to Root changed, it/they will advertise the new mapping, and the traffic will re-converge to the new primary path. MP then withdraw the backup label after finishing their re-converge. Pn will delete the specified backup LSP like as the process of normally P2MP LSP. And the entire backup P2MP LSP will be deleted when all the node MP leave the backup P2MP LSP.

5.1.2. PLR Switching Over Considerations

The P2MP Based Node Protection also has the BFD scalability issue on the Protected node. Similar with P2P Based Node Protection solution, this section provides two methods for deployment.

- Option 1:
  If PLR can not differentiate link and node failure, MP must take the responsibility to drop one of the two reduplicate traffic 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 detected and MP will drop traffic on the backup path until it sees N fails.

- 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 failure. In this case, the Node Failure Required Flag, in the P2MP Based MP Node Protection Status Element, must be set as 'Y'.

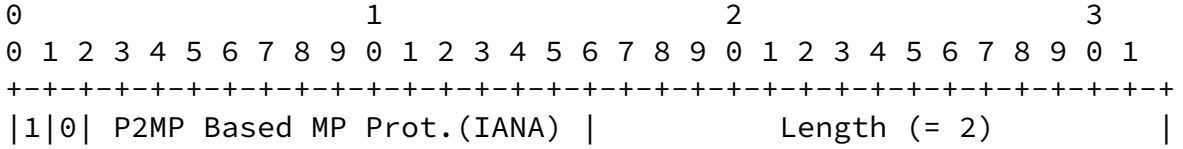
Note that, all the MPs of N MUST use one same Node Failure Required Flag value. Otherwise, the backup P2MP LSP tree need depart to two trees different from the switch over type, and this part is TBD. And it is also possible that can use a backup MP2MP LSP tree to protect one node in the primary MP2MP LSP tree, this part is TBD too

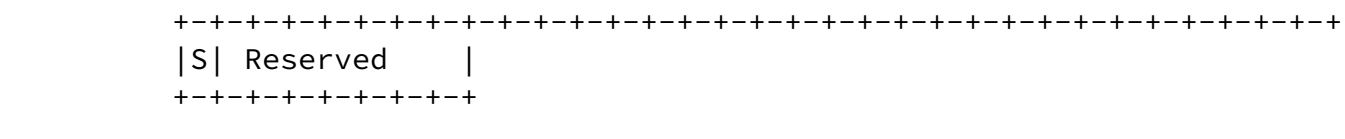
[Editors Note - This Editors note and remaining options will be removed before publication of this document.]

5.2. Protocol Extensions for P2MP Based Node Protection

5.2.1. P2MP Based MP Protection Capability Parameter TLV

A new Capability Parameter TLV is defined as P2MP Based MP Protection Capability for node protection. Following is the format of this new Capability Parameter TLV:





S: As specified in [RFC5561]

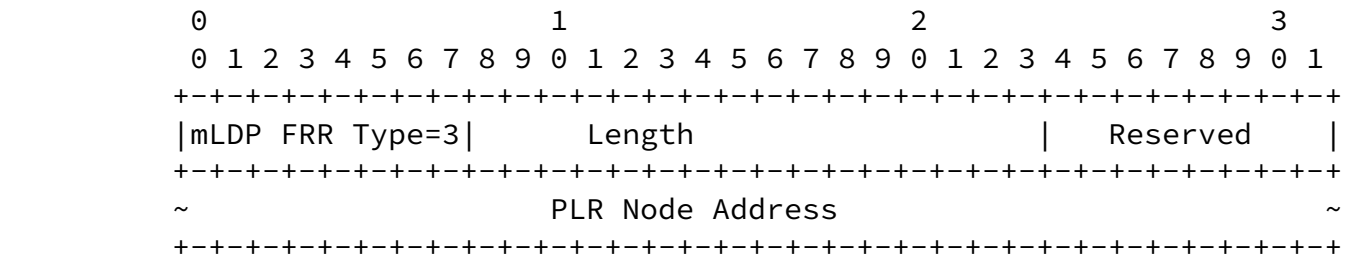
This is an unidirectional capability announced.

An LSR, which supports the P2MP based protection procedures, should advertise this P2MP Based MP Protection Capability TLV to its LDP speakers. Without receiving this capability annoucement, an LSR MUST NOT send any message including the P2MP Based MP Node Protection Status 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. This part is TBD.

5.2.2. P2MP Based MP Node Protection Status Elements

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

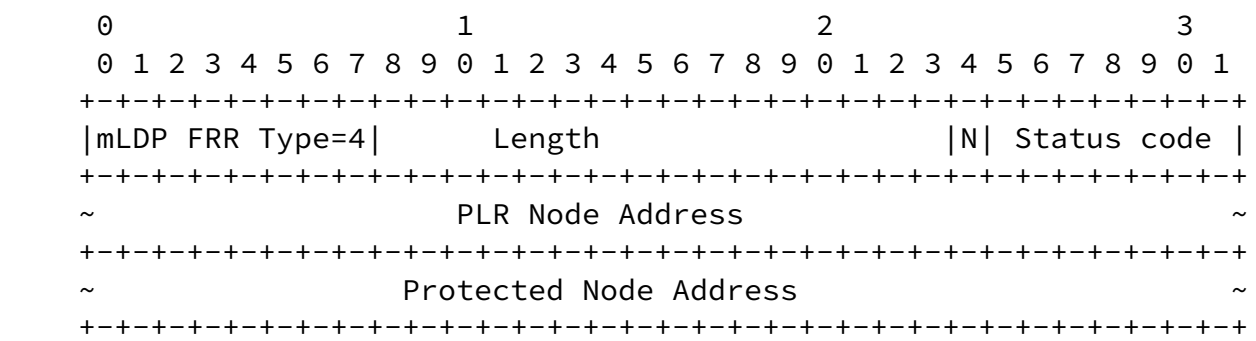


mLDP FRR Type: Type 3 (to be assigned by IANA)

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.

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



mLDP FRR Type: Type 4 (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

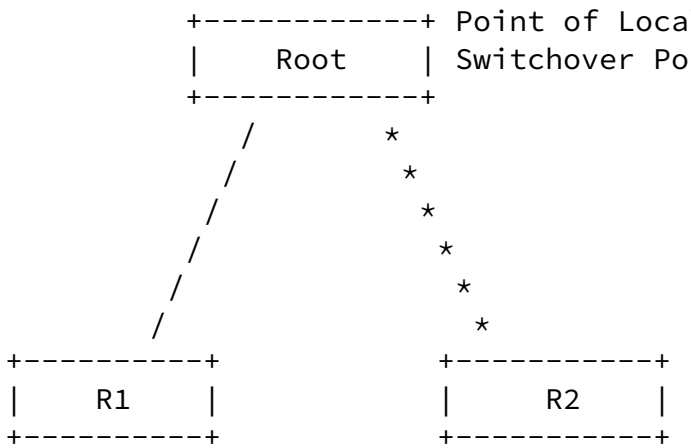
PLR Node Address: The host address of the PLR Node.

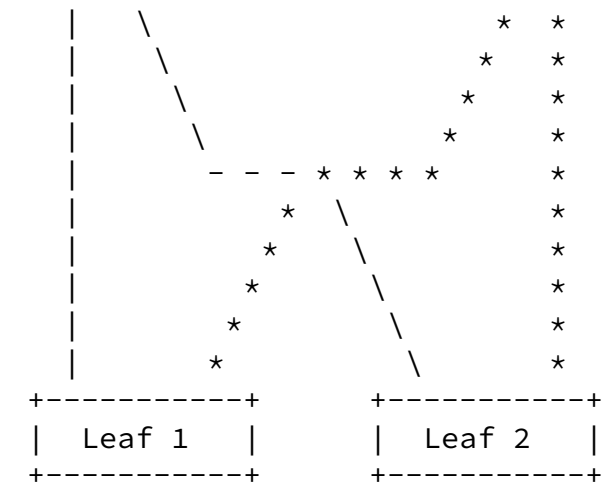
Protected Node Address: The host address of the Protected Node.

N Bit: Node Failure Required Flag, which indicates the swichover timing 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.

6. mLDP End-to-End Protection using LDP/mLDP Multiple Topology

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

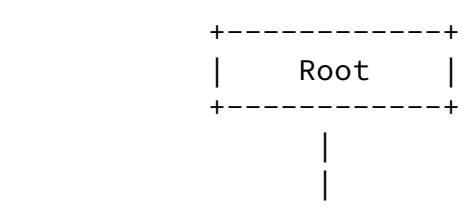


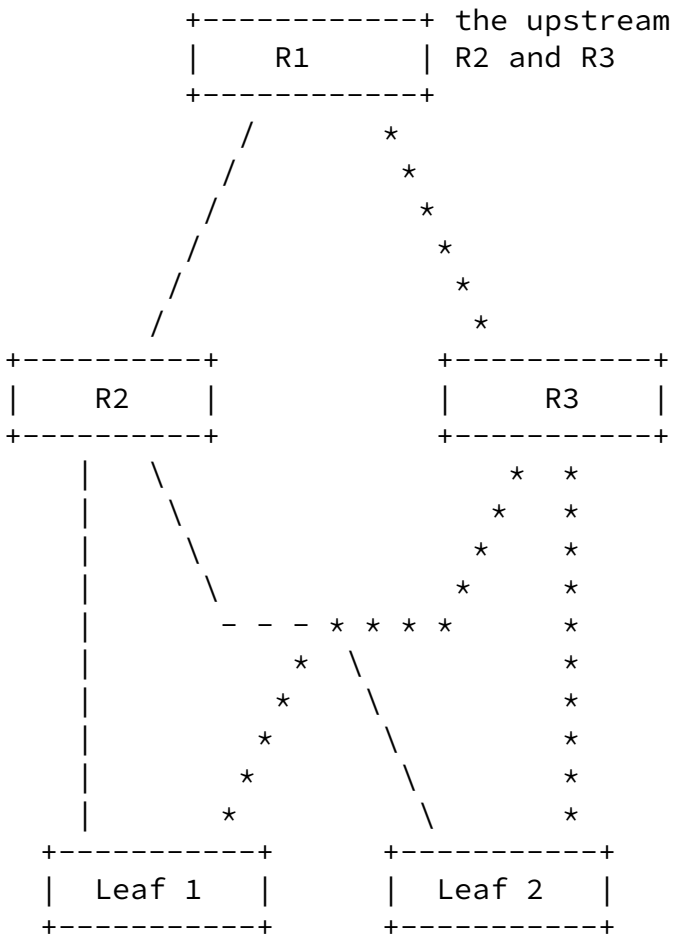


mLDP End-to-end Protection Example

Figure 3

In Figure 3 (mLDP End-to-end Protection Example), there are two separated topologies from Root node to Leaf 1 and Leaf 2. For the same root address and opaque value, 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.





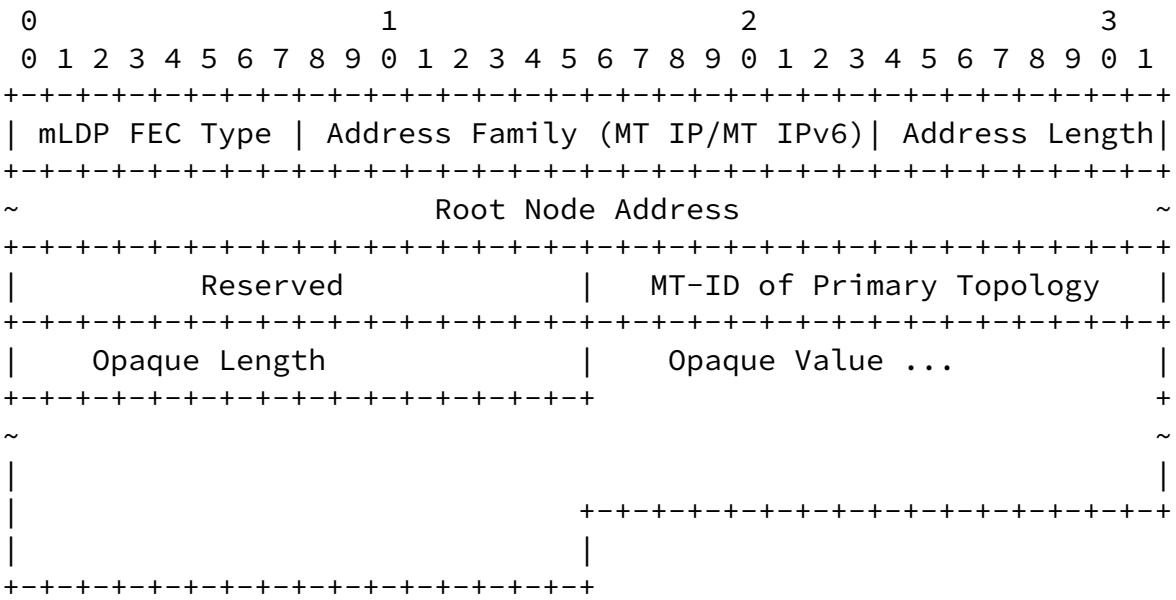
mLDP End-to-end Protection with Shared Upstream Node

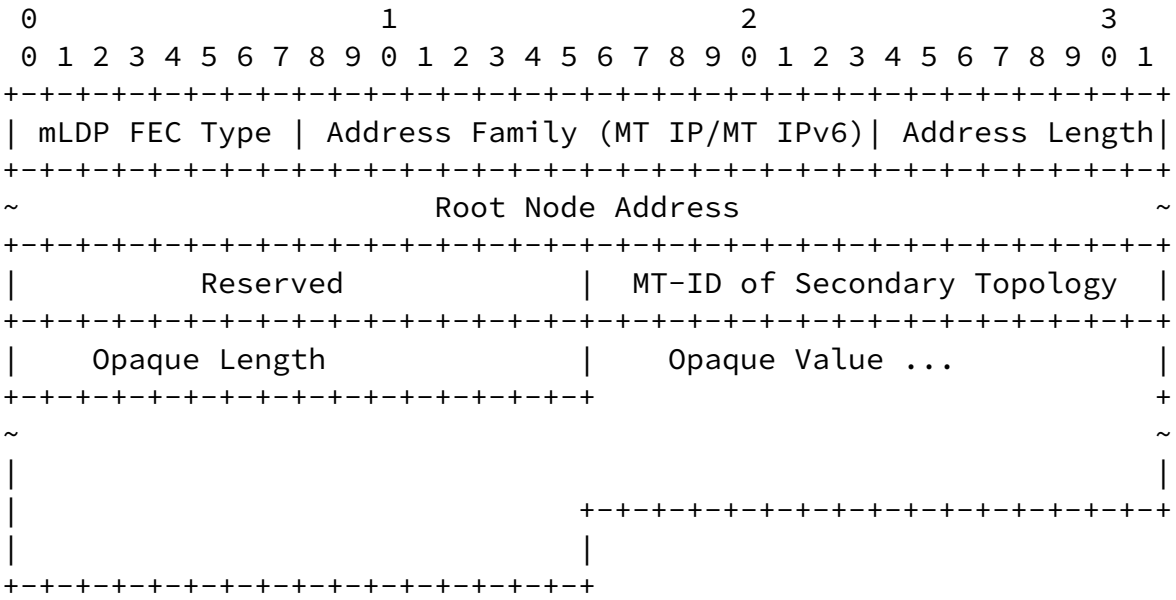
Figure 4

In Figure 4 (mLDP End-to-end Protection with Shared Upstream Node Example), there are two separated topologies from Root node to Leaf 1 and Leaf 2 except the link between R1 and Root node. For the same root address and opaque value, 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. The difference in this example comparing to the last example where the primary and backup topology are totally disjoint, if there is a link failure between the Root and R1 or node R1 fails, there is no protection available.

6.1. Signaling Procedures for MT Based End-to-end Protection

Using the protocol extensions and signaling procedure provided by [I-D.ietf-mpls-ldp-multi-topology], Leaf 1 and Leaf 2 in figure 8 or figure 9 are able to trigger mLDP LSPs in different topologies, sending label mapping messages with same root address, same opaque value, different MT-ID and different label. Based on the two new Address Families named "MT IP" and "MT IPv6" introduced in [I-D.ietf-mpls-ldp-multi-topology] that can be used to specify IP prefixes within a topology scope, the mLDP FEC elements for leaf1 and leaf2 will be encoded as follows:





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. Either Root node or Leaf node can be the Failure Detector.

6.2. Protocol extensions for MT Based End-to-end Protection

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

7. IANA Considerations

This memo includes the following requests to IANA:

- P2P Based MP Protection Capability.
- P2MP Based MP Protection Capability.
- mLDP P2P Encapsulation type for LDP MP Status Value Element.
- mLDP FRR types for LDP MP Status Value Element.

8. Manageability Considerations

[Editors Note - This section requires further discussion]

8.1. Control of Function and Policy

8.2. Information and Data Models

8.3. Liveness Detection and Monitoring

8.4. Verifying Correct Operation

8.5. Requirements on Other Protocols and Functional Component

8.6. Impact on Network Operation

8.7. Policy Control

9. 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.

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

10. 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. We also would like to thank Robin Li, Lujun Wan, Alia Atlas and IJsbrand Wijnands for their comments and suggestions to the draft.

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.

[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.

11.2. Informative References

[I-D.ietf-mpls-ldp-multi-topology]
Zhao, Q., Fang, L., Zhou, C., Li, L., and N. So, "LDP Extensions for Multi Topology Routing", draft-ietf-mpls-ldp-multi-topology-03 (work in progress), March 2012.

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

Tao Chou
Huawei Technology
156 Beiqing Rd
Beijing, P.R. 100095
China

Email: tao.chou@huawei.com

Daniel King
Old Dog Consulting

Email: E-mail: daniel@olddog.co.uk