

MPLS Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: February 17, 2022

L. Han  
China Mobile  
F. Yang  
Huawei Technologies  
J. Zhao  
CAICT  
August 16, 2021

Signal Degrade Indication in Segment Routing over MPLS Network  
draft-han-mpls-sdi-sr-02

Abstract

This document describes a typical use case of MPLS-TP, where signal degrade defect needs to be correctly detected and transmitted via OAM messages within network. When MPLS-TP evolves to Segment Routing MPLS, transit node has no knowledge of labels to be encapsulated in MPLS label stack. Transit node cannot spread OAM messages with signal degrade defect indication. Thus, a solution is proposed in this draft.

Requirements 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].

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 <https://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 February 17, 2022.

## Copyright Notice

Copyright (c) 2021 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 (<https://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.

## Table of Contents

1. Background . . . . .	2
2. Terminology . . . . .	3
3. Problem Statement . . . . .	4
3.1. Defect Triggered Procedure . . . . .	4
3.2. MPLS-TP Solution . . . . .	4
3.3. Problem in SR-MPLS . . . . .	6
4. Solution in SR-MPLS . . . . .	6
5. IANA Considerations . . . . .	7
6. Security Considerations . . . . .	7
7. Acknowledgements . . . . .	7
8. References . . . . .	7
8.1. Normative References . . . . .	8
8.2. Informative References . . . . .	8
Authors' Addresses . . . . .	8

## 1. Background

In early era of telecommunication, transport network is set up to provide voice service. The connection in network is always connection-oriented and circuit switching. With the rapid increasing bandwidth brought by Ethernet, transport network transforms into the packet-switched transport network. Technologies like MPLS/PWE3 perfectly meet the requirements of supporting both packet-transport and circuit-transport. It led to the work of MPLS Transport Profile (MPLS-TP), collaborated between ITU-T and IETF at the first decade of the 21st century.

MPLS-TP is a subset of MPLS. Features that are not applicable to transport network are excluded, and features to meet the requirements of transport network, e.g., bidirectional path, deterministic control and management, etc., are strictly required. According to the Joint

Working Team consensus, any extension of MPLS-TP would be included in MPLS field.

With the emerge of Segment Routing (SR) and Software Defined Network (SDN), MPLS-TP network technologies are adapted as well. In this draft, we recognize one use case where the signal degrade defect can be correctly detected and transmitted via MPLS-TP OAM in MPLS-TP, but not fulfilled in SR-MPLS. To fix this problem is the motivation of this draft.

Editor's note: This section gives a historical introduction of MPLS-TP, since it has been extensively deployed in packet switched transport networks for years. The intention of this section is to help readers understand the unique of requirements from packet transport network. Once the draft becomes RFC, part of this section can be moved to Appendix.

## 2. Terminology

MPLS: MultiProtocol Label Switching

PWE3: Pseudo Wire Emulation Edge to Edge

MPLS-TP: MultiProtocol Label Switching - Transport Profile

SR: Segment Routing

SDN: Software Defined Network

OAM: Operation, Administration and Maintenance

SD: Signal Degrade

BER: Bit Error Rate

WDM: Wavelength Division Multiplexing

NMS: Network Management System

G-ACh: Generic Associated Channel

PDU: Protocol Data Unit

CCM: Continuity Check Message

MEP: Maintenance Entity Group End Point

MIP: Maintenance Entity Group Intermediate Point

AIS: Alarm Indication Signal

### 3. Problem Statement

#### 3.1. Defect Triggered Procedure

Signal Degrade (SD) describes a status of signal associated data has degraded and a degraded defect is active. Signal degrade of a physical link is usually measured and represented by Bit Error Rate (BER) value. Fiber aging, impairment and pollution, optical module mismatch or WDM transmission error are the reasons to lead to signal degrade. More information about signal degrade can be found in [I-D.yang-mpls-ps-sdi-sr].

In practice, when physical link degrades in network, signal degrade defect is firstly detected and reported by the node. A specific type of alarm is generated and sent to Network Management System (NMS) or a SDN controller. It is a report to management plane and strongly required from perspective of network management. However, the problem is the notification to management plane is usually not fast enough to assist the network recovery. It may result in hour or even day level of service interruption time.

As mentioned in [RFC6372], defect may trigger system to perform a survivability action, when notification of an issue is reported from equipment in a lower layer, system fails to receive an OAM continuity check message, or receives of an OAM message reporting a failure condition. Similarly, when signal degrade defect is reported from the lower layer, e.g. physical layer, local protection mechanism can be triggered within the internal system of nodez. In case of protection switchover selector is at the source or destination node, while the signal degrade is happened at intermediate node, an OAM message should be transmitted to notify the degrade condition to the nodes actually perform the protection switchover. This action is preferred to be triggered by events in the data plane [RFC6372].

#### 3.2. MPLS-TP Solution

Generic Associated Channel (G-ACh) [RFC5586] is defined to carry OAM messages for MPLS pseudowires, LSPs and sections. The Generic Associated Channel format used in MPLS is shown in Figure 1. By using the generic associated channel and indication of channel type, different OAM mechanisms with different formats can be encapsulated uniformly as well as independently.

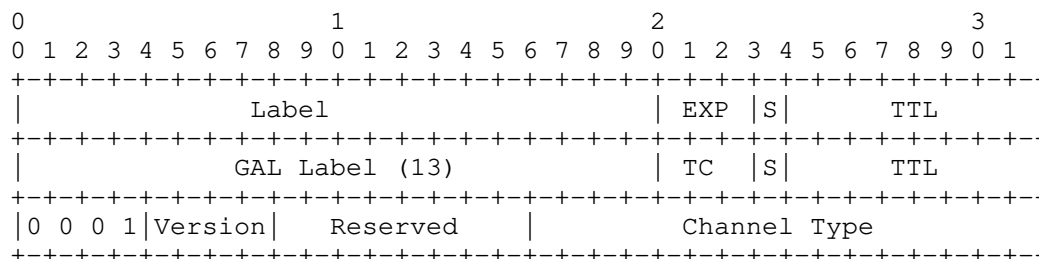


Figure 1 G-ACh Format in MPLS

In MPLS-TP, ITU-T G.8113.1 [ITU-T G.8113.1] specifies a large set of OAM mechanisms and has been widely deployed in packet transport networks. Figure 2 shows the common OAM PDU format of different OAM mechanisms.

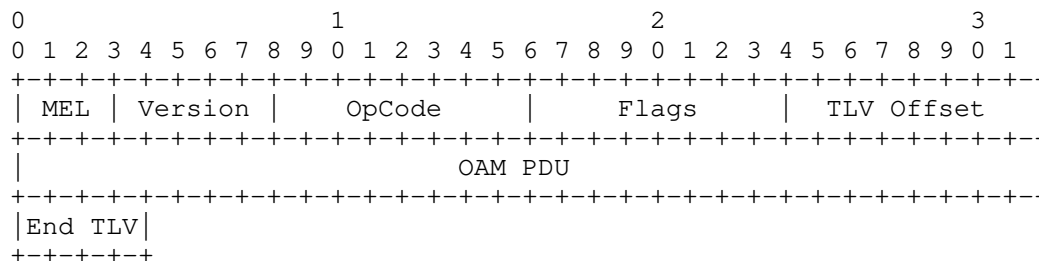


Figure 2 ITU-T G.8113.1 Common OAM PDU Format

**MEG Level:** MEG Level is a 3-bit field. It contains an integer value that identifies the MEG level of OAM PDU. Value ranges from 0 to 7.

**Version:** Version is a 5-bit field. It contains an integer value that identifies the OAM protocol version. Value is 0 in the current version.

**OpCode:** OpCode is a 1-octet field. It contains an OpCode that identifies an OAM PDU type. OpCode is used to identify the remaining content of an OAM PDU. Value for the CCM PDU type is 1.

**Flags:** Flags is an 8-bit field. Use of the bits in this field is dependent on the OAM PDU type.

**TLV Offset:** TLV Offset is a 1-octet field. It contains the offset to the first TLV in an OAM PDU relative to the TLV Offset field. The value of this field is associated with an OAM PDU type. When the TLV Offset is 0, it points to the first octet following the TLV Offset field.

End TLV: an all-ZEROes octet value.

When signal degrade happens in MPLS-TP, an MPLS-TP Alarm Indication Signal (AIS) OAM message with active AIS indication is generated and transmitted within the OAM maintenance domain. Maintenance Entity Group End Point (MEP), usually also acting as protection switchover selector, performs the protection switchover once it receives the AIS indication in MPLS-TP OAM message.

### 3.3. Problem in SR-MPLS

When Segment Routing is introduced to MPLS, the nodes except the headend have no information of the forwarding path. If the signal degrade is happened on the transit nodes, MPLS-TP AIS OAM message cannot be generated because this node has no knowledge of labels ought to be encapsulated in MPLS label stack. Either the label information of forwarding path can be obtained on transit node, or the defect can be indicated in different messages could help the defect spread in network. It is valuable to keep transit node with the capability of reporting defects in SR-MPLS.

### 4. Solution in SR-MPLS

Segment routing is designed to reduce the states in transit nodes, any defects like SD defect cannot be indicated in a newly generated OAM message on transit node. Alternative way is to indicate the defect in other OAM messages. Continuity Check Message (CCM) is proposed to indicate the signal degrade defect for two reasons. Firstly, CCM is designed to be applicable for fault management, performance monitoring, or protection switching applications. Secondly, consider the merit of CCM's various transmission period, the defect indication can be flexibly transmitted according to operator's needs.

One reservation bits in Flag section in CCM OAM PDU message can be used as Error Indication (EI) to indicate signal degrade. Flag format with EI extension is shown in Figure 3.

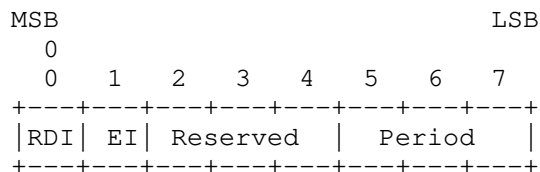


Figure 3 CCM OAM PDU Flags Format with EI Extension

RDI: Remote Defect Indication, set to 1 to indicate RDI, otherwise it is set to 0.

Period: Indicate the transmission period.

EI: Error indication, 0 indicates no error, 1 indicates error.

Reserved: Reserved fields are set to all ZEROes.

If the node detects the signal degrade defect, EI field is set in CCM OAM message and transmitted to other nodes. Note that, Maintenance Entity Group Intermediate Point (MIP) is required to be transparent to CCM message in MPLS-TP. In order to support BER indication on each node along the forwarding path, extra configuration and intervening implementation to process CCM message would be required on MIP.

Editor's Note: When other OAM mechanisms used in generic associated channel (G-ACh), there might be various solutions to transmit signal degrade defect, or any other defects detected by transit nodes. This draft introduces a very light-weight solution, which has already been implemented and deployed in networks.

## 5. IANA Considerations

This document requests IANA to assign one bit from Flags of MPLS-TP OAM PDU format to indicate "Signal Degrade".

## 6. Security Considerations

There are MEP and MIP node defined in OAM mechanisms. Some types of OAM message are defined to be transparent to MIP node, and requires no extra configuration or message processing on MIP nodes. If the transit node of SR-MPLS acts as MIP in OAM maintenance domain, this MIP node needs to process the OAM messages to indicate the defects. At the moment, explicit configuration is required on MIP to have the authority to process OAM messages.

## 7. Acknowledgements

The authors want to thank Yuanlong Jiang, Mach Chen, Yongjian Hu for their valuable suggestions during the construction of draft.

## 8. References

## 8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

## 8.2. Informative References

- [I-D.yang-mpls-ps-sdi-sr]  
Yang, F., Han, L., and J. Zhao, "Problem Statement of Signal Degrade Indication for SR over MPLS", draft-yang-mpls-ps-sdi-sr-01 (work in progress), November 2020.
- [RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed., "MPLS Generic Associated Channel", RFC 5586, DOI 10.17487/RFC5586, June 2009, <<https://www.rfc-editor.org/info/rfc5586>>.
- [RFC6372] Sprecher, N., Ed. and A. Farrel, Ed., "MPLS Transport Profile (MPLS-TP) Survivability Framework", RFC 6372, DOI 10.17487/RFC6372, September 2011, <<https://www.rfc-editor.org/info/rfc6372>>.
- [ITU-T\_G8113.1]  
ITU-T, "ITU-T G.8113.1: Operations, administration and maintenance mechanisms for MPLS-TP in packet transport networks", April 2016.

## Authors' Addresses

Liuyan Han  
China Mobile  
Beijing  
China

Email: [hanliuyan@chinamobile.com](mailto:hanliuyan@chinamobile.com)

Fan Yang  
Huawei Technologies  
Beijing  
China

Email: [shirley.yangfan@huawei.com](mailto:shirley.yangfan@huawei.com)



Junfeng Zhao  
CAICT  
Beijing  
China

Email: zhaojunfeng@caict.ac.cn

MPLS Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: 16 June 2022

K. Kompella  
R. Balaji  
Juniper Networks  
R. Thomas  
Cohesity  
13 December 2021

Label Distribution Using ARP  
draft-kompella-mpls-larp-11

Abstract

This document describes extensions to the Address Resolution Protocol to distribute MPLS labels for IPv4 and IPv6 host addresses. Distribution of labels via ARP enables simple plug-and-play operation of MPLS, which is key to deploying MPLS in data centers and enterprises.

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 <https://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 16 June 2022.

Copyright Notice

Copyright (c) 2021 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 (<https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

1. Introduction . . . . .	2
1.1. Requirements Language . . . . .	3
1.2. Approach . . . . .	3
2. Overview of Ethernet ARP . . . . .	3
3. L-ARP Protocol Operation . . . . .	4
3.1. Setup . . . . .	5
3.2. Egress Operation . . . . .	5
3.3. Ingress Operation . . . . .	6
3.4. Data Plane . . . . .	6
4. Attributes . . . . .	7
4.1. Secondary Attributes . . . . .	7
5. L-ARP Message Format . . . . .	8
5.1. Hardware Address Format . . . . .	9
5.2. CT TLV . . . . .	9
6. Security Considerations . . . . .	10
7. IANA Considerations . . . . .	10
8. Acknowledgments . . . . .	10
9. References . . . . .	10
9.1. Normative References . . . . .	10
9.2. Informative References . . . . .	11
Authors' Addresses . . . . .	12

## 1. Introduction

This document describes extensions to the Address Resolution Protocol (ARP) [RFC0826] to advertise label bindings for IP host addresses. While there are well-established protocols, such as LDP [RFC5036], RSVP [RFC3209], BGP [RFC3107] and SPRING-MPLS [RFC8660], that provide robust mechanisms for label distribution, these protocols tend to be relatively complex, and often require detailed configuration for proper operation. There are situations where a simpler protocol may be more suitable from an operational standpoint. An example is the case where an MPLS Fabric is the underlay technology in a Data Center; here, MPLS tunnels originate from host machines. The host thus needs a mechanism to acquire label bindings to participate in the MPLS Fabric, but in a simple, plug-and-play manner. Existing signaling/routing protocols do not always meet this need. Labeled

ARP (L-ARP) is a proposal to fill that gap.

### 1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

The term "server" will be used in this document to refer to an ARP/L-ARP server; the term "host" will be used to refer to a compute server or other device acting as an ARP/L-ARP client.

### 1.2. Approach

ARP is a nearly ubiquitous protocol; every device with an Ethernet interface, from hand-helds to hosts, have an implementation of ARP. ARP is plug-and-play; ARP clients do not need configuration to use ARP. That suggests that ARP may be a good fit for devices that want to source and sink MPLS tunnels, but do so in a zero-config, plug-and-play manner, with minimal impact to their code.

The approach taken here is to create a minor variant of the ARP protocol, labeled ARP (L-ARP), which is distinguished by a new hardware type, MPLS-over-Ethernet. Regular (Ethernet) ARP (E-ARP) and L-ARP can coexist; a device, as an ARP client, can choose to send out an E-ARP or an L-ARP request, depending on whether it needs Ethernet or MPLS connectivity. Another device may choose to function as an E-ARP server and/or an L-ARP server, depending on its ability to provide an IP-to-Ethernet and/or IP-to-MPLS mapping.

## 2. Overview of Ethernet ARP

In the most straightforward mode of operation [RFC0826], ARP queries are sent to resolve "directly connected" IP addresses. The ARP request is broadcast, with the Target Protocol Address field (see Section 5 for a description of the fields in an ARP message) carrying the IP address of another node in the same subnet. All the nodes in the LAN receive this ARP request. All the nodes, except the node that owns the IP address, ignore the ARP request. The IP address owner learns the MAC address of the sender from the Source Hardware Address field in the ARP request, and unicasts an ARP reply to the sender. The ARP reply carries the replying node's MAC address in the Source Hardware Address field, thus enabling two-way communication between the two nodes.

A variation of this scheme, known as "proxy ARP" [RFC2002], allows a node to respond to an ARP request with its own MAC address, even when the responding node does not own the requested IP address. Generally, the proxy ARP response is generated by routers to attract traffic for prefixes they can forward packets to. This scheme requires the host to send ARP queries for the IP address the host is trying to reach, rather than the IP address of the router. When there is more than one router connected to a network, proxy ARP enables a host to automatically select an exit router without running any routing protocol to determine IP reachability. Unlike regular ARP, a proxy ARP request can elicit multiple responses, e.g., when more than one router has connectivity to the address being resolved. The sender must be prepared to select one of the responding routers.

Yet another variation of the ARP protocol, called 'Gratuitous ARP' [RFC2002], allows a node to update the ARP cache of other nodes in an unsolicited fashion. Gratuitous ARP is sent as either an ARP request or an ARP reply. In either case, the Source Protocol Address and Target Protocol Address contain the sender's address, and the Source Hardware Address is set to the sender's hardware address. In case of a gratuitous ARP reply, the Target Hardware Address is also set to the sender's address.

### 3. L-ARP Protocol Operation

The L-ARP protocol builds on the proxy ARP model, and also leverages gratuitous ARP model for asynchronous updates.

In this memo, we will refer to L-ARP clients (that make L-ARP requests) and L-ARP servers (that send L-ARP responses). In Figure 1, H1, H2 and H3 are L-ARP clients, and T1, T2 and T3 are L-ARP servers. T4 is a member of the MPLS Fabric that may not be an L-ARP server. Within the MPLS Fabric, the usual MPLS protocols (IGP (i.e., SPRING-MPLS), LDP, RSVP-TE) are run. Say H1, H2 and H3 want to establish MPLS tunnels to each other (for example, they are using BGP MPLS VPNs as the overlay virtual network technology). H1 might also want to talk to a member of the MPLS Fabric, say T. Also, the "protocol" addresses in L-ARP requests are either IPv4 or IPv6 addresses; note that while it is common to use Neighbor Discovery (ND) [RFC4861] for "regular" ARP requests when dealing with IPv6 (i.e., to obtain Ethernet MAC addresses corresponding to an IPv6 address), ND is not used when the ARP request is for an MPLS label.

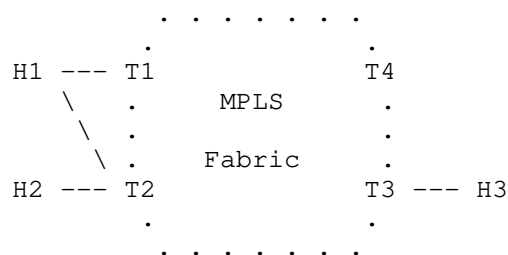


Figure 1: MPLS Fabric

### 3.1. Setup

In Figure 1, the nodes T1-T4, and those in between making up the "MPLS Fabric" are assumed to be running some protocol whereby they can signal MPLS reachability to themselves and to other nodes (like hosts H1-H3). T1-T3 are L-ARP servers; T4 need not be, since it doesn't have an attached L-ARP client. H1-H3 are L-ARP clients.

### 3.2. Egress Operation

A node (say T3) that wants an attached node (say H3) to have MPLS reachability allocates a label L3 to reach H3 and advertises this label into the MPLS Fabric. This can be triggered by configuration on T3, or when T3 first receives an L-ARP request from H3 (indicating that H3 wants MPLS connectivity), or via some other protocol. T3 then advertises (H3, L3) to its peers in the MPLS Fabric so that all members of the Fabric have connectivity to H3. This advertisement can be one of the following:

- \* a "proxy" LDP message (sent on behalf of H3) with prefix H3 and label L3; or
- \* a node Segment ID (SID) advertised on behalf of H3; or
- \* a labeled BGP advertisement, with prefix H3, label L3 and next hop self.

On receiving a packet with label L3, T3 pops the label and send the packet to H3. (In the case of labeled BGP, there would be a two-label stack, with outer label to reach T3 and inner label of L3.) This is the usual operation of an MPLS Fabric, with the addition of advertising labels for nodes outside the fabric.

### 3.3. Ingress Operation

A node (say H1, an L-ARP client) that needs an MPLS tunnel to another node (say H3) identified by a host address (either IPv4 or IPv6) broadcasts over all its interfaces an L-ARP request with the Target Protocol Address set to H3 and Hardware Type set to "MPLS-over-Ethernet". A node receiving the L-ARP request (say T1, an L-ARP server) does the following:

1. checks if it has MPLS reachability to H3. If not, it ignores the L-ARP request.
2. if it does, T1 allocates a label TL3 to reach H3 (if it doesn't already have such a label) and installs an L-FIB entry to swap L1 with the label (stack) to reach H3.
3. sends a (proxy) L-ARP reply to H1 with the Source Hardware Address (SHA) set to (L, M), where M is T1's metric to H3. T1 may also set some attribute bits in the SHA.

### 3.4. Data Plane

To send a packet to H3 over an MPLS tunnel, H1 pushes L1 onto the packet, sets the destination MAC address to M1 and sends it to T1. On receiving this packet, T1 swaps the top label with the label(s) for its MPLS tunnel to H3. If T1's reachability to H3 is via a SPRING label stack, the label L1 acts as an implicit binding SID.

If H1 and H3 have an overlay connection (say an IPVPN [RFC4364] VPN-foo) whereby VM1 on H1 wishes to talk to VM3 on H3 over VPN-foo, H1 does the following:

1. H1 learns information about VPN-foo via BGP (or an SDN controller), including the VPN label VL3 to use to talk to VM3;
2. H1 installs a VRF for VPN-foo, with prefix VM3, label VL3 and next hop H3;
3. H1 binds the local "veth" interface to VM1 to this VRF.
4. When VM1 sends a packet to dest IP address VM3 over its veth interface, H1 looks up VM3 in the corresponding VRF, gets label VL3. It then sends an L-ARP request for next hop H3, and gets TL3.
5. Finally, H1 pushes the label pair (TL3, VL3) onto the packet from VM1 and sends this to T1. This packet will then end up at VM3 on H3.

Note that H1 broadcasts its L-ARP request over its attached interfaces. H1 may receive several L-ARP replies; in that case, H1 can select any subset of these to send MPLS packets destined to H3. As described later, the L-ARP response may contain certain parameters that enable the client to make an informed choice. If the target H3 belongs to one of the subnets that H1 participates in, and H3 is capable of sending L-ARP replies, H1 can use H3's response to send MPLS packets to H3.

#### 4. Attributes

In addition to carrying a label stack to be used in the data plane, an L-ARP reply carries some attributes that are typically used in the control plane. One of these is a metric. The metric is the distance from the L-ARP server to the destination. This allows an L-ARP client that receives multiple responses to decide which ones to use, and whether to load-balance across some of them. The metric typically will be the IGP shortest path distance from server to the destination; this makes comparing metrics from different servers meaningful.

Another attribute is Entropy Label (EL) Capability. This attribute says whether the destination is EL capable (ELC). In Figure 1, if T3 advertises a label to reach H3 and T3 is ELC, T3 can include in its signaling to T1 that it is ELC. In that case, T1's L-ARP reply to H1 can have ELC bit set. This tells H1 that it may include (below the outermost label) an Entropy Label Indicator followed by an Entropy Label. This will help improve load balancing across the MPLS Fabric, and possibly on the last hop to H3.

##### 4.1. Secondary Attributes

Beyond the basic attributes that are carried with every L-ARP request, there are more optional attributes, for example, to ask for certain characteristics of the path traffic takes to the destination. These attributes are carried in TLVs that are carried in L-ARP requests and replies.

One such TLV is the Classful Transport (CT: see [I-D.kaliraj-idr-bgp-classful-transport-planes]) TLV. This TLV allows the L-ARP client to request a label to a destination over a tunnel of the given Transport Class. To satisfy this request, the L-ARP server creates (or finds) a tunnel to the destination that is routed over the CT Transport Plane, allocates a label L, inserts an entry in the LFIB to swap L to the tunnel, and sends L to the L-ARP client in its reply.



## 5. L-ARP Message Format

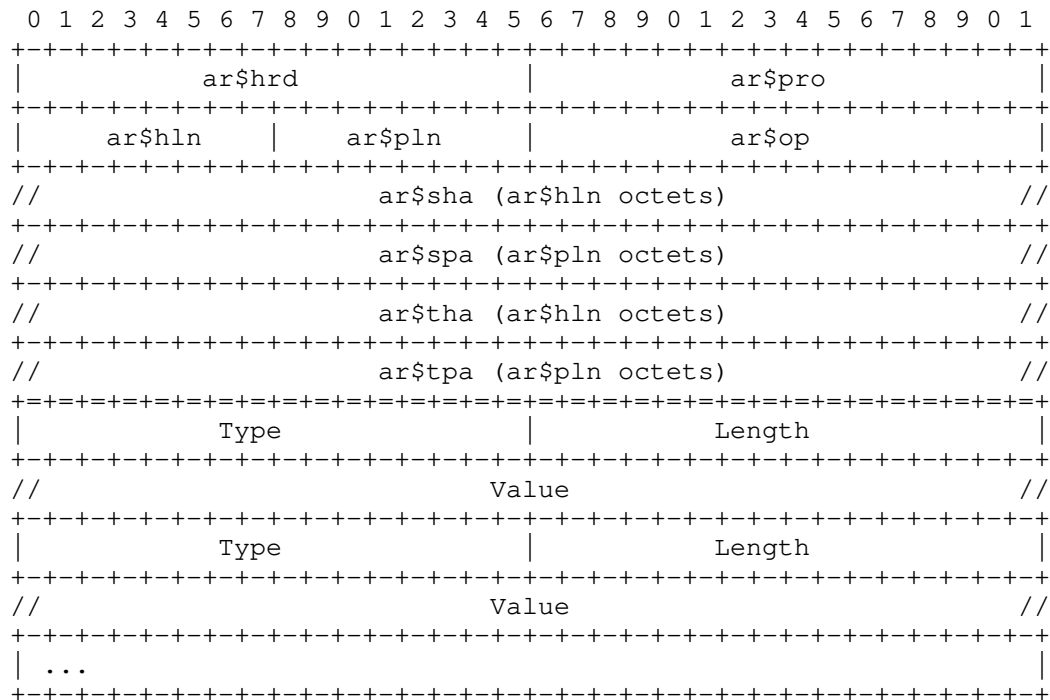


Figure 2: L-ARP Packet Format

ar\$hrd: Hardware Type: MPLS-over-Ethernet. The value of the field used here is HTYPE-MPLS. To start with, we will use the experimental value HW\_EXP2 (256).

ar\$pro: Protocol Type: IPv4/IPv6. The value of the field used here is 0x0800 to resolve an IPv4 address and 0x86DD to resolve an IPv6 address.

ar\$hln: Hardware Address Length: 6

ar\$pln: Protocol Address Length: for an IPv4 address, the length is 4 octets; for an IPv6 address, it is 16.

ar\$op: Operation Code: set to 1 for request, 2 for reply, and 10 for ARP-NAK. Other op codes may be used as needed.

ar\$sha: Source Hardware Address: In an L-ARP request, this is usually all zeros. In an L-ARP reply, Source Hardware Address is the label to reach ar\$spa, as specified in Figure 3 below.

ar\$spa: Source Protocol Address: In an L-ARP request, this field carries the sender's IP address. In an L-ARP reply, this field carries the requested IP address (which may not be the sender's IP address).

ar\$tha: Target Hardware Address: In an L-ARP message, this is all zeros.

ar\$tpa: Target Protocol Address: In an L-ARP request, this field carries the IP address for which the client is seeking an MPLS label.

Type: a 2-octet field defining the Type of the TVL

Length: a 2-octet field defining the Length L of the TVL

Value: an L-octet field with the Value of the TLV

### 5.1. Hardware Address Format

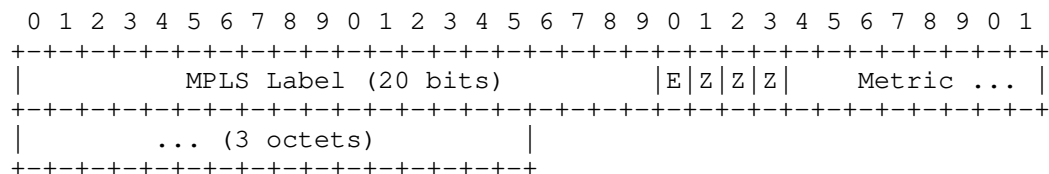


Figure 3: Label Format in L-ARP

MPLS Label: The 20-bit label

E-bit: Entropy Label Capable: this flag indicates whether the corresponding label in the label stack can be followed by an Entropy Label. If this flag is set, the client has the option of inserting ELI and EL as specified in [RFC6790]. The client can choose not to insert ELI/EL pair. If this flag is clear, the client must not insert ELI/EL after the corresponding label.

Z: These bits are not used, and SHOULD be set to zero on sending and ignored on receipt.

Metric: The 3-octet IGP metric to ar\$tha from the point of view of the L-ARP replier.

### 5.2. CT TLV

The CT TLV has Type (TBD) and Length 4 octets; the Value field consists of the CT attribute.

## 6. Security Considerations

There are many possible attacks on ARP: ARP spoofing, ARP cache poisoning and ARP poison routing, to name a few. These attacks use gratuitous ARP as the underlying mechanism, a mechanism used by L-ARP. Thus, these types of attacks are applicable to L-ARP. Furthermore, ARP does not have built-in security mechanisms; defenses rely on means external to the protocol.

It is well outside the scope of this document to present a general solution to the ARP security problem. One simple answer is to add a TLV that contains a digital signature of the contents of the ARP message. This TLV would be defined for use only in L-ARP messages, although in principle, other ARP messages could use it as well. Such an approach would, of course, need a review and approval by the Security Directorate. If approved, the type of this TLV and its procedures would be defined in this document. If some other technique is suggested, the authors would be happy to include the relevant text in this document, and refer to some other document for the full solution.

## 7. IANA Considerations

IANA is requested to allocate a new ARP hardware type (from registry hrd) for HTYPE-MPLS [IANA].

IANA is further requested to create a registry for Types of L-ARP Secondary Attributes. This registry should contain an entry for the CT Type Section 5.2.

## 8. Acknowledgments

Many thanks to Shane Amante for his detailed comments and suggestions. Many thanks to the team in Juniper prototyping this work for their suggestions on making this variant workable in the context of existing ARP implementations. Thanks too to Luyuan Fang, Alex Semenyaka and Dmitry Afanasiev for their comments and encouragement.

## 9. References

### 9.1. Normative References

- [IANA] IANA, "Address Resolution Protocol (ARP) Parameters/ Hardware Types", n.d., <<https://www.iana.org/assignments/arp-parameters/arp-parameters.xhtml>>.

- [RFC0826] Plummer, D., "An Ethernet Address Resolution Protocol: Or Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, RFC 826, DOI 10.17487/RFC0826, November 1982, <<https://www.rfc-editor.org/info/rfc826>>.
- [RFC2002] Perkins, C., Ed., "IP Mobility Support", RFC 2002, DOI 10.17487/RFC2002, October 1996, <<https://www.rfc-editor.org/info/rfc2002>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <<https://www.rfc-editor.org/info/rfc6790>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

## 9.2. Informative References

- [I-D.kaliraj-idr-bgp-classful-transport-planes] Vairavakkalai, K., Venkataraman, N., Rajagopalan, B., Mishra, G., Khaddam, M., Xu, X., Szarecki, R. J., and D. J. Gowda, "BGP Classful Transport Planes", Work in Progress, Internet-Draft, draft-kaliraj-idr-bgp-classful-transport-planes-12, 25 August 2021, <<https://www.ietf.org/archive/id/draft-kaliraj-idr-bgp-classful-transport-planes-12.txt>>.
- [RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", RFC 3107, DOI 10.17487/RFC3107, May 2001, <<https://www.rfc-editor.org/info/rfc3107>>.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <<https://www.rfc-editor.org/info/rfc3209>>.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <<https://www.rfc-editor.org/info/rfc4364>>.

- [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <<https://www.rfc-editor.org/info/rfc4861>>.
- [RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <<https://www.rfc-editor.org/info/rfc5036>>.
- [RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <<https://www.rfc-editor.org/info/rfc8660>>.

## Authors' Addresses

Kireeti Kompella  
Juniper Networks  
1133 Innovation Way  
Sunnyvale, 94089  
United States of America

Phone: +1-408-745-2000  
Email: [kireeti.ietf@gmail.com](mailto:kireeti.ietf@gmail.com)

Balaji Rajagopalan  
Juniper Networks  
Survey No.111/1 to 115/4, Wing A & B  
Bangalore 560103  
India

Email: [balajir@juniper.net](mailto:balajir@juniper.net)

Reji Thomas  
Cohesity

Email: [rejithomas.d@gmail.com](mailto:rejithomas.d@gmail.com)

MPLS WG  
Internet-Draft  
Intended status: Standards Track  
Expires: January 13, 2022

K. Kompella  
W. Lin  
Juniper Networks  
July 12, 2021

No Further Fast Reroute  
draft-kompella-mpls-nffrr-02

## Abstract

There are several cases where, once Fast Reroute has taken place (for MPLS protection), a second fast reroute is undesirable, even detrimental. This memo gives several examples of this, and proposes a mechanism to prevent further fast reroutes.

## 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 <https://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 13, 2022.

## Copyright Notice

Copyright (c) 2021 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 (<https://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.

## Table of Contents

1. Introduction . . . . .	2
1.1. Other Approaches . . . . .	3
1.2. Terminology . . . . .	3
2. Motivation . . . . .	3
2.1. EVPN (VPN/VPLS) Active-active Multihoming . . . . .	3
2.2. RMR Protection . . . . .	4
2.3. General MPLS forwarding . . . . .	5
3. Solution . . . . .	6
3.1. NFFRR for MPLS forwarding . . . . .	6
3.2. Proposal . . . . .	8
3.2.1. NFFRR and SPRING . . . . .	10
3.3. NFFRR for MPLS Services . . . . .	10
3.4. NFFRR for RMR . . . . .	11
4. Signaling NFFRR Capability . . . . .	12
4.1. Signaling NFFRR Capability for MPLS Services with BGP . . . . .	12
4.2. Signaling NFFRR Capability for MPLS Services with Targeted LDP . . . . .	12
4.3. Signaling NFFRR Capability for MPLS Forwarding . . . . .	12
5. IANA Considerations . . . . .	12
6. Security Considerations . . . . .	13
7. References . . . . .	13
7.1. Normative References . . . . .	13
7.2. Informative References . . . . .	14
Authors' Addresses . . . . .	15

## 1. Introduction

MPLS Fast Reroute (FRR) [RFC4090] [RFC5286] [RFC7490] is a useful and widely deployed tool for minimizing packet loss in the case of a link or node failure. This has not only proven to be very effective, it is often the reason for using MPLS as a data plane. FRR works for a variety of control plane protocols, including LDP, RSVP-TE, and SPRING. Furthermore, FRR is often used to protect MPLS services such as IP VPN and EVPN.

Having said this, there are case where, once FRR has taken place, if the packet encounters a second failure, a second FRR is not helpful, perhaps even disruptive. For example, the packet may loop until TTL expires. This can lead to link congestion and further packet loss. Thus, the attempt to prevent a packet from being dropped may instead affect many other packets. Note that the "second" failure may simply be another manifestation of the same failure; see Figure 1.

This memo proposes a mechanism for preventing further FRR once in cases where such further protection may be harmful. Several examples where this is the case are demonstrated as motivation. A solution

using special-purpose labels (SPLs) is then offered. Some mechanisms for distributing the capability to avoid further fast reroutes are also discussed, although these may be better placed in other documents in other Working Groups.

### 1.1. Other Approaches

[ALDT] has a more elaborate mechanism for preventing loops due to multiple failures. This involves marking the nodes redirecting traffic in a header (either individually, or as node groups), and dropping the packet at a transit node if its ID is in the header.

### 1.2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

## 2. Motivation

A few cases are given where "further fast reroute" is harmful. Some of the cases are for MPLS services; others for "plain" MPLS forwarding.

### 2.1. EVPN (VPN/VPLS) Active-active Multihoming

Consider the following topology for multihoming an Ethernet VPN (EVPN [RFC7432]) Customer Edge (CE) device for protection against the failure of a Provider Edge (PE) device or a PE-CE link. To do so, there is a backup MPLS path between PE2 and PE3 (denoted by the starred line).

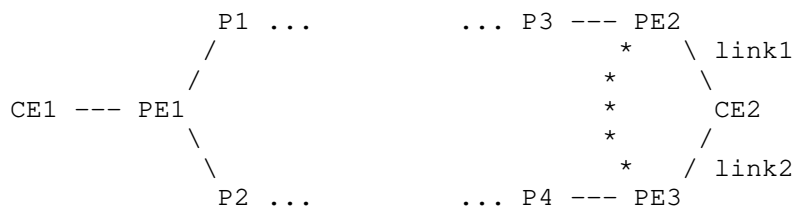


Figure 1: EVPN Multihoming

Suppose (known unicast) traffic goes from CE1 to CE2. With active-active multihoming, this traffic will be load-balanced between PE2 (to CE2 via link link1) and PE3 (to CE2 via link2). If link1 were to



fail, PE2 can still get traffic for CE2 by sending it over the backup path to PE3 (and similarly for PE3 if link2 fails).

However, suppose CE2 is down. PE2 will assume link1 is down and send traffic for CE2 to PE3 over the backup path. PE3 (which thinks that link2 is down; note that the single real failure of CE2 being down is manifested as separate failures to PE2 and PE3) will protect this "second" failure by sending traffic for CE2 over the backup path to PE2. Thus, traffic will ping-pong between PE2 and PE3 until TTL expires.

Thus, the attempt to protect traffic to CE2 may end up doing more harm than good, by congesting the backup path between PE2 and PE3 and by giving PE2 and PE3 useless work to do.

A similar topology can be used in EVPN-Etree [RFC8317], EVPN-VPWS [RFC8214], IP VPN [RFC4364] or VPLS [RFC4761] [RFC4762]. In all these cases, the same looping behavior would occur for unicast traffic if CE2 is down.

## 2.2. RMR Protection

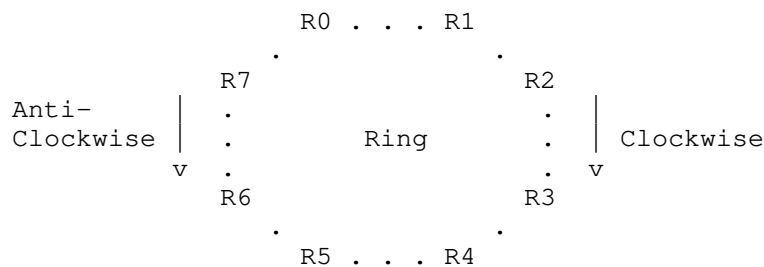


Figure 2: RMR Looping

In Resilient MPLS Rings (RMR), suppose traffic goes from a node, say R0, to a node, say R4, over a clockwise path. Protection consists of switching this traffic onto the anti-clockwise path to R4. This works well if a node or link between R0 or R4 is down. However, if node R4 itself is down, its adjacent neighbor R3, will send the traffic anti-clockwise to R4; when this traffic reaches R4's other neighbor R5, it will return to N3, and so on, until TTL expires. [I-D.ietf-mpls-rmr] provides more details, and offers some means of mitigation. This memo offers a more elegant solution.

### 2.3. General MPLS forwarding

Consider the following topology:

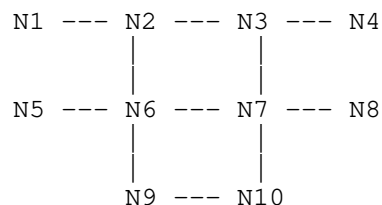


Figure 3: General MPLS Forwarding

Say link protection is configured for links N2-N3 and N6-N7. Link N2-N3 is protected by a bypass tunnel N2-N6-N7-N3, and link N7-N3 is protected by a bypass tunnel N7-N6-N2-N3. (These bypass tunnels may be set up using RSVP-TE [RFC3209] or via SPRING stacks [RFC8660].) Say furthermore that there is an LSP from N1 to N4 with path N1-N2-N3-N4, which asks for link protection. If link N2-N3 fails, traffic will take the path N1-N2-N6-N7-N3-N4.

Suppose, however, links N2-N3 and N7-N3 fail simultaneously. This may happen if they share fate (e.g., go over a common fiber conduit); it may also appear to happen if node N3 fails. Either way, first, the bypass protecting link N2-N3 kicks in, and traffic is sent to N3 via N6 and N7. However, when the traffic hits N7, the bypass for N7-N3 kicks in, and traffic is sent back to N2. Thus the traffic will loop between N2 and N7 until TTL expires, in the process congesting links N2-N6 and N6-N7.

Now consider an LSP: N5-N6-N7-N8. The link N6-N7 may be protected by the bypass N6-N2-N3-N7 or by N6-N9-N10-N7, or by load-balancing between these two bypasses. If both links N2-N3 and N6-N7 fail, then traffic that is protected via bypass N6-N2-N3-N7 will ping-pong between N6 and N2 until TTL expires; traffic protected via bypass N6-N9-N10-N7 will successfully make it to N8. If link N6-N7 is protected by load-balancing across the two bypass paths, then about half the traffic will loop between N6 and N2, and the rest will make it to N8.

While the above description is for protection using a bypass tunnel, the same principle applies to protection using Loop-Free Alternates [RFC5286] [RFC7490] or any of its variants (such as Topology Independent LFA).

### 3. Solution

To address this issue, we suggest the use of a SPL [RFC7274] called NFFRR (value TBD; suggested: 8). An alternate would be to use an extended SPL, whereby a pair of labels indicates that no further fast route is desired. However, in the case of SPRING MPLS bypass tunnels (Section 3.2.1) of depth N, this would triple the label stack size. Using regular SPLs instead would only double the stack size.

#### 3.1. NFFRR for MPLS forwarding

To illustrate, we'll first take the example of Figure 3, with MPLS paths signaled using RSVP-TE. This method can be used for paths that use SPRING stacks, but this will be detailed in a later version.

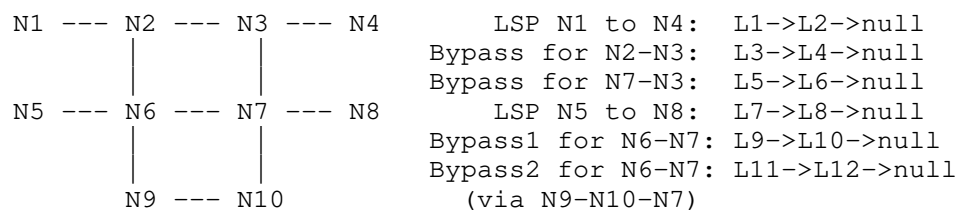


Figure 4: Example Using RSVP-TE LSPs

Node	Action	Next	New Pkt	Comment
N1	push L1	N2	[L1] pkt	ingress
N2	L1 -> L2	N3	[L2] pkt	
N3	pop L2	N4	pkt	PHP
N4	fwd pkt	-	-	continue

Table 1: Forwarding from N1 to N4

Note 1: "[L1 ...]" denotes the label stack on the packet; pkt is the original packet received at ingress. "L1 -> L2" means swap label L1 with L2. "pop L2" means pop the top label L2. "fwd pkt" means forward the packet as usual.

Node	Action	Next	New Pkt	Comment
N2	push L3	N6	[L3] pkt	ingress
N6	L3 -> L4	N7	[L4] pkt	
N7	pop L4	N3	pkt	PHP

Table 2: Forwarding over the bypass for link N2-N3

Node	Action	Next	New Pkt	Comment
N7	push L5	N6	[L5] pkt	ingress
N6	L5 -> L6	N2	[L6] pkt	
N2	pop L6	N3	pkt	PHP

Table 3: Forwarding over Bypass1 for link N7-N3

Node	Action	Next	New Pkt	Comment
N1	push L1	N2	[L1] pkt	ingress
N2	L1 -> L2	N3	[L2] pkt	N3 X
N2	push L3	N6	[L3 L2] pkt	PLR
N6	L3 -> L4	N7	[L4 L2] pkt	
N7	pop L4	N3	[L2] pkt	merge
N3	pop L2	N4	pkt	PHP
N4	fwd pkt	-	-	continue

Table 4: Forwarding from N1 to N4 if link N2-N3 fails

Table 4 is obtained by composing Table 1 and Table 2.

Note 2: "N3 X" means "next hop N3 unavailable (because link N2-N3 failed)".

Node	Action	Next	New Pkt	Comment
N1	push L1	N2	[L1] pkt	ingress
N2	L1 -> L2	N3	[L2] pkt	N3 X
N2	push L3	N6	[L3 L2] pkt	PLR
N6	L3 -> L4	N7	[L4 L2] pkt	
N7	pop L4	N3	[L2] pkt	N3 X'
N7	push L5	N6	[L5 L2] pkt	
N6	L5 -> L6	N2	[L6 L2] pkt	PLR
N2	pop L6	N3	[L2] pkt	N3 X
N2	push L3	N6	[L3 L2]	PLR
etc				loop!

Table 5: Forwarding from N1 to N4 if links N2-N3 and N7-N3 fail

Table 5 is obtained by composing Table 1, Table 2 and Table 3.

Note 3: "N3 X'" means "next hop N3 unavailable because link N7-N3 is down.

Note 4: While the impact of a loop is pretty bad, the impact of an ever-growing label stack (not illustrated here) and possible associated fragmentation on transit nodes may be worse.

### 3.2. Proposal

An LSR (typically a PLR) that wishes to prevent further FRRs after the first one can push an SPL, namely NFFRR, onto the label stack as follows:

Node	Action	Next	New Pkt	Comment
N1	push L1	N2	[L1] pkt	ingress
N2	L1 -> L2	N3	[L2] pkt	N3 X
N2	push L3, NFFRR	N6	[L3 NFFRR L2] pkt	PLR
N6	L3 -> L4	N7	[L4 NFFRR L2] pkt	
N7	pop L4, NFFRR	N3	[L2] pkt	merge
N3	pop L2	N4	pkt	PHP
N4	fwd pkt	-	-	continue

Table 6: Forwarding from N1 to N4 if link N2-N3 fails with NFFRR

Note 5: N2 can insert an NFFRR label only if it knows that all LSRs in the path can process it correctly. See Section 4 for some details on how this capability is communicated.

Node	Action	Next	New Pkt	Comment
N1	push L1	N2	[L1] pkt	ingress
N2	L1 -> L2	N3	[L2] pkt	N3 X
N2	push L3, NFFRR	N6	[L3 NFFRR L2] pkt	PLR
N6	L3 -> L4	N7	[L4 NFFRR L2] pkt	
N7	pop L4	N3	[NFFRR L2] pkt	N3 X
N7	check NFFRR	-	-	drop pkt

Table 7: Forwarding from N1 to N4 if links N2-N3 and N7-N3 fail with NFFRR

Note 6: "check NFFRR" means that, before N7 applies FRR (because link N7-N3 is down), N7 checks the label below the top label (or in this case, because of PHP, the top label itself). If this is the NFFRR label, N7 drops the packet rather than apply FRR.

### 3.2.1. NFFRR and SPRING

Suppose that, to protect link N2-N3, a bypass tunnel N2-N6-N7-N3 were instantiated using SPRING MPLS [RFC8660], in particular, using adjacency SIDs. If the corresponding labels for links N6-N7 and N7-N3 were L20 and L21, the bypass would consist of pushing the label stack [L20 L21] onto the packet and sending the packet to N6. To indicate that FRR has already occurred and to drop the packet rather than to try to protect the packet again, N2 would have to push [L20 NFFRR L21 NFFRR] onto the packet before sending it to N6. If the packet came from N1 with label L1, N2 would send a packet with label stack [L20 NFFRR L21 NFFRR L2] to N6.

N6 would see L20, pop it, note the NFFRR label and pop it, then attempt to send the packet to N7. If the link N6-N7 is down, N6 drops the packet. Otherwise, N7 gets the packet, sees L21, pops it, sees NFFRR, pops it and tries to send the packet to N3. If link N7-N3 is down, N7 drops the packet. Otherwise, N3 gets the packet with L2, swaps with with L3 and sends it to N4.

Note that with SPRING MPLS, the NFFRR label needs to be repeated for each label in the bypass stack. Hence the request for a "regular" SPL rather than an extended SPL.

### 3.3. NFFRR for MPLS Services

First, we illustrate known unicast EVPN forwarding:

Node	Action	Next	Packet	Comment
PE1	send to CE2	PE2	[T1 S2] pkt	EVPN
PE2	send to CE2	link1	pkt	done!

Note: T1/T2/T3 are the transport labels for PE1/PE3/PE2 to reach PE2/PE2/PE3 respectively. S2/S3 are the service labels announced by PE2/PE3 for CE2.

Then, we show what happens when CE2 is down without NFFRR:

Node	Action	Next	Packet	Comment
PE1	send to CE2	PE2	[T1 S2] pkt	EVPN
PE2	send to CE2	link1	--	link1 X
PE2	send to CE2	PE3	[T3 S3] pkt	eFRR
PE3	send to CE2	link2	--	link2 X
PE3	send to CE2	PE2	[T2 S2] pkt	eFRR
PE2	send to CE2	link1	--	link1 X
PE2	send to CE2	PE3	[T3 S3] pkt	eFRR
...				loop!

Note: link1/link2 X means link1/link2 is down. eFRR refers to EVPN multihoming FRR.

In the case of MPLS services such as EVPN Figure 1, the NFFRR label is inserted below the service label, as shown below:

Node	Action	Next	Packet	Comment
PE1	send to CE2	PE2	[T1 S2] pkt	EVPN
PE2	send to CE2	link1	--	link1 X
PE2	send to CE2	PE3	[T3 S2 NFFRR] pkt	eFRR
PE3	send to CE2	link2	--	link2 X
PE3	drop pkt	--	--	check NFFRR

Note: "check NFFRR" is as above.

### 3.4. NFFRR for RMR

As described in Figure 2, packets will loop until TTL expires if the destination node in an RMR ring (here, R4) fails. The solution in this case is that the first node to apply RMR protection (R3) pops the current RMR transport label being used, sees that the next label



is not NFFRR (so protection is allowed), pushes an NFFRR label and then the RMR transport label for the reverse direction.

When R5 receives the packet, it sees that the next link is down, pops the RMR transport label, sees the NFFRR label and drops the packet. Thus, the loop is avoided.

#### 4. Signaling NFFRR Capability

##### 4.1. Signaling NFFRR Capability for MPLS Services with BGP

The ideal choice would be an attribute consisting of a bit vector of node capabilities, one bit of which would be the capability of processing the NFFRR SPL below the BGP service label. This would be used by BGP L2VPN, BGP VPLS, EVPN, E-Tree and E-VPWS. An alternative is to use the BGP Capabilities Optional Parameter [I-D.ietf-idr-next-hop-capability]. Details to be worked out.

##### 4.2. Signaling NFFRR Capability for MPLS Services with Targeted LDP

One approach to signaling NFFRR capability for MPLS services signaled with targeted LDP is to introduce a new LDP TLV called the NFFRR Capability TLV as an Optional Parameter in the Label Mapping Message [RFC5036]. This TLV has Type TBD (suggested: 0x0207) and Length 0.

Another approach is to use LDP Capabilities [RFC5561]; this approach has the advantage that it deals with capabilities on a node basis rather than on a per label mapping basis. However, there don't appear to be other documents using this approach.

##### 4.3. Signaling NFFRR Capability for MPLS Forwarding

The authors suggest signaling a router's ability to process the NFFRR SPL using the Link State Router TE Node Capabilities [RFC5073], which works for both IS-IS and OSPF. A new TE Node Capability bit, the N bit (suggested value 5) indicates that the advertising node is capable of processing the NFFRR SPL.

#### 5. IANA Considerations

If this draft is deemed useful, a way to signal that No Further Fast-route should be performed on a packet will be needed. There are two approaches: allocate an SPL for NFFRR: if so, we suggest the early allocation of label 8 for this. Alternatively, if [I-D.kompella-mpls-mspl4fa] (or similar) is adopted, allocate a forwarding action bit saying whether or not to do FRR.

Furthermore, means of signaling the ability to process the NFFRR SPL/bit should be defined for IS-IS, OSPF, LDP and BGP.

The following update is suggested for the Link State Router TE Node Capabilities registry:

Bit	Name	Reference
5	NFFRR	This document

The following update is suggested for the TLV Type Name Space of the Label Distribution Protocol (LDP) Parameters registry:

Type	Name	Reference
0x0207	NFFRR	This document

## 6. Security Considerations

A malicious or compromised LSR can insert NFFRR into a label stack, preventing FRR from occurring. If so, protection will not kick in for failures that could have been protected, and there will be unnecessary packet loss.

## 7. References

### 7.1. Normative References

- [I-D.kompella-mppls-mspl4fa]  
Kompella, K., Beeram, V. P., Saad, T., and I. Meilik,  
"Multi-purpose Special Purpose Label for Forwarding  
Actions", draft-kompella-mppls-mspl4fa-00 (work in  
progress), February 2021.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate  
Requirement Levels", BCP 14, RFC 2119,  
DOI 10.17487/RFC2119, March 1997,  
<<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed.,  
"LDP Specification", RFC 5036, DOI 10.17487/RFC5036,  
October 2007, <<https://www.rfc-editor.org/info/rfc5036>>.

- [RFC5073] Vasseur, J., Ed. and J. Le Roux, Ed., "IGP Routing Protocol Extensions for Discovery of Traffic Engineering Node Capabilities", RFC 5073, DOI 10.17487/RFC5073, December 2007, <<https://www.rfc-editor.org/info/rfc5073>>.
- [RFC7274] Kompella, K., Andersson, L., and A. Farrel, "Allocating and Retiring Special-Purpose MPLS Labels", RFC 7274, DOI 10.17487/RFC7274, June 2014, <<https://www.rfc-editor.org/info/rfc7274>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

## 7.2. Informative References

- [ALDT] Merling, D., Braun, W., and M. Menth, "Efficient Data Plane Protection for SDN", June 2018, <<https://atlas.informatik.uni-tuebingen.de/~menth/papers/Menth18g.pdf>>.
- [I-D.ietf-idr-next-hop-capability] Decraene, B., Kompella, K., and W. Henderickx, "BGP Next-Hop dependent capabilities", draft-ietf-idr-next-hop-capability-06 (work in progress), October 2020.
- [I-D.ietf-mpls-rmr] Kompella, K. and L. M. Contreras, "Resilient MPLS Rings", draft-ietf-mpls-rmr-14 (work in progress), February 2021.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <<https://www.rfc-editor.org/info/rfc3209>>.
- [RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, DOI 10.17487/RFC4090, May 2005, <<https://www.rfc-editor.org/info/rfc4090>>.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <<https://www.rfc-editor.org/info/rfc4364>>.
- [RFC4761] Kompella, K., Ed. and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, DOI 10.17487/RFC4761, January 2007, <<https://www.rfc-editor.org/info/rfc4761>>.

- [RFC4762] Lasserre, M., Ed. and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, DOI 10.17487/RFC4762, January 2007, <<https://www.rfc-editor.org/info/rfc4762>>.
- [RFC5286] Atlas, A., Ed. and A. Zinin, Ed., "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, DOI 10.17487/RFC5286, September 2008, <<https://www.rfc-editor.org/info/rfc5286>>.
- [RFC5561] Thomas, B., Raza, K., Aggarwal, S., Aggarwal, R., and JL. Le Roux, "LDP Capabilities", RFC 5561, DOI 10.17487/RFC5561, July 2009, <<https://www.rfc-editor.org/info/rfc5561>>.
- [RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <<https://www.rfc-editor.org/info/rfc7432>>.
- [RFC7490] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", RFC 7490, DOI 10.17487/RFC7490, April 2015, <<https://www.rfc-editor.org/info/rfc7490>>.
- [RFC8214] Boutros, S., Sajassi, A., Salam, S., Drake, J., and J. Rabadan, "Virtual Private Wire Service Support in Ethernet VPN", RFC 8214, DOI 10.17487/RFC8214, August 2017, <<https://www.rfc-editor.org/info/rfc8214>>.
- [RFC8317] Sajassi, A., Ed., Salam, S., Drake, J., Uttaro, J., Boutros, S., and J. Rabadan, "Ethernet-Tree (E-Tree) Support in Ethernet VPN (EVPN) and Provider Backbone Bridging EVPN (PBB-EVPN)", RFC 8317, DOI 10.17487/RFC8317, January 2018, <<https://www.rfc-editor.org/info/rfc8317>>.
- [RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <<https://www.rfc-editor.org/info/rfc8660>>.

Authors' Addresses

Kireeti Kompella  
Juniper Networks  
1133 Innovation Way  
Sunnyvale, CA 94089  
United States

Email: kireeti.kompella@gmail.com

Wen Lin  
Juniper Networks  
1133 Innovation Way  
Sunnyvale, CA 94089  
United States

Email: wlin@juniper.net

MPLS Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: 8 September 2022

IJ. Wijnands  
Individual  
K. Raza  
M. Mishra  
A. Budhiraja  
Cisco Systems, Inc.  
Z. Zhang  
Juniper Networks  
A. Gulko  
Edward Jones wealth management  
7 March 2022

mLDP Extensions for Multi-Topology Routing  
draft-wijnands-mpls-mlbp-multi-topology-04

Abstract

Multi-Topology Routing (MTR) is a technology to enable service differentiation within an IP network. Flexible Algorithm (FA) is another mechanism of creating a sub-topology within a topology using defined topology constraints and computation algorithm. In order to deploy mLDP in a network that supports MTR and/or FA, mLDP is required to become topology and FA aware. This document specifies extensions to mLDP to support MTR with FA such that when building a Multi-Point LSPs it can follow a particular topology and algorithm.

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 <https://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 8 September 2022.

Copyright Notice

Copyright (c) 2022 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 (<https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

1. Glossary . . . . .	2
2. Introduction . . . . .	3
3. Specification of Requirements . . . . .	4
4. MT Scoped mLDP FECs . . . . .	4
4.1. MP FEC Extensions for MT . . . . .	4
4.1.1. MP FEC Element . . . . .	5
4.1.2. MT IP Address Families . . . . .	5
4.1.3. MT MP FEC Element . . . . .	6
4.2. Topology IDs . . . . .	7
5. MT Multipoint Capability . . . . .	8
6. MT Applicability on FEC-based features . . . . .	9
6.1. Typed Wildcard MP FEC Elements . . . . .	9
6.2. End-of-LIB . . . . .	10
7. Topology-Scoped Signaling and Forwarding . . . . .	10
7.1. Upstream LSR selection . . . . .	10
7.2. Downstream forwarding interface selection . . . . .	10
8. LSP Ping Extensions . . . . .	10
9. Security Considerations . . . . .	11
10. IANA Considerations . . . . .	11
11. Acknowledgments . . . . .	12
12. References . . . . .	12
12.1. Normative References . . . . .	12
12.2. Informative References . . . . .	13
Authors' Addresses . . . . .	13

## 1. Glossary

MT - Multi-Topology

MT-ID - Multi-Topology Identifier

MTR - Multi-Topology Routing

IGP - Interior Gateway Protocol

MP - Multipoint (P2MP or MP2MP)

LDP - Label Distribution Protocol

mLDP - Multipoint LDP

P2MP - Point-to-Multipoint

MP2MP - Multipoint-to-Multipoint

FEC - Forwarding Equivalence Class

LSP - Label Switched Path

FA - Flexible Algorithm

IPA - IGP Algorithm

## 2. Introduction

Multi-Topology Routing (MTR) is a technology to enable service differentiation within an IP network. IGP protocols (OSPF and IS-IS) and LDP have already been extended to support MTR. To support MTR, an IGP maintains independent IP topologies, termed as "Multi-Topologies" (MT), and computes/installs routes per topology. OSPF extensions [RFC4915] and ISIS extensions [RFC5120] specify the MT extensions under respective IGPs. To support IGP MT, similar LDP extensions [RFC7307] have been specified to make LDP MT-aware and be able to setup unicast Label Switched Paths (LSPs) along IGP MT routing paths.

A more light weight mechanism to define constraint-based topologies is Flexible Algorithm (FA) [I-D.ietf-lsr-flex-algo]. FA can be seen as creating a sub-topology within a topology using defined topology constraints and computation algorithm. This can be done within a MTR topology or just the default Topology. An instance of such a sub-topology is identified by a 1 octet value as documented in [I-D.ietf-lsr-flex-algo]). Flexible Algorithm is a mechanism to create a sub-topology, but in the future different algorithms might be defined on how to achieve that. For that reason, in the remainder of this document we'll refer to this as the IGP Algorithm (IPA).



Multipoint LDP (mLDP) refers to extensions in LDP to setup multipoint LSPs (point-to-multipoint (P2MP) or multipoint-to-multipoint (MP2MP)), by means of set of extensions and procedures defined in [RFC6388]. In order to deploy mLDP in a network that supports MTR and FA, mLDP is required to become topology and algorithm aware. This document specifies extensions to mLDP to support MTR/IPA such that when building a Multi-Point LSPs it can follow a particular topology and algorithm. This means that the identifier for the particular Topology to be used by mLDP have to become a two tuple (MTR Topology Id, IGP Algorithm).

### 3. Specification of Requirements

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 [RFC2119].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

### 4. MT Scoped mLDP FECs

As defined in [RFC7307], MPLS Multi-Topology Identifier (MT-ID) is an identifier that is used to associate an LSP with a certain MTR topology. In the context of MP LSPs, this identifier is part of the mLDP FEC encoding so that LDP peers are able to setup an MP LSP via their own defined MTR policy. In order to avoid conflicting MTR policies for the same mLDP FEC, the MT-ID needs to be a part of the FEC, so that different MT-ID values will result in unique MP-LSP FEC elements.

The same applies to the IPA. The IPA needs to be encoded as part of the mLDP FEC to create unique MP-LSPs and at the same time is used to signal to mLDP (hop-by-hop) which Algorithm needs to be used to create the MP-LSP.

Since the MT-ID and IPA are part of the FEC, they apply to all the LDP messages that potentially include an mLDP FEC element.

#### 4.1. MP FEC Extensions for MT

Following subsections propose the extensions to bind an mLDP FEC to a topology. The mLDP MT extensions reuse some of the extensions specified in [RFC7307].

## 4.1.1. MP FEC Element

Base mLDP specification [RFC6388] defines MP FEC Element as follows:

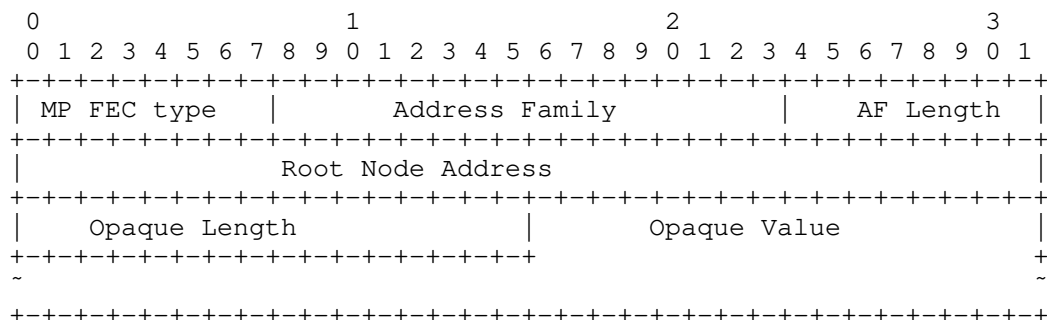


Figure 1: MP FEC Element Format [RFC6388]

Where "Root Node Address" encoding is as defined for given "Address Family", and whose length (in octets) is specified by the "AF Length" field.

To extend MP FEC elements for MT, the {MT-ID, IPA} is a tuple that is relevant in the context of the root address of the MP LSP. The {MT-ID, IPA} tuple determines in which (sub)-topology the root address needs to be resolved. Since the {MT-ID, IPA} tuple should be considered part of the mLDP FEC, the most natural place to encode this tuple is as part of the root address. While encoding it, we also propose to use "MT IP" Address Families as described in following sub section.

## 4.1.2. MT IP Address Families

[RFC7307] has specified new address families, named "MT IP" and "MT IPv6", to allow specification of an IP prefix within a topology scope. In addition to using this address family for mLDP, we also use 8 bits of the 16 bits Reserved field to encode the IGP Algorithm (IPA) Registry. The resulting format of the data associated with these new Address Families is as follows:

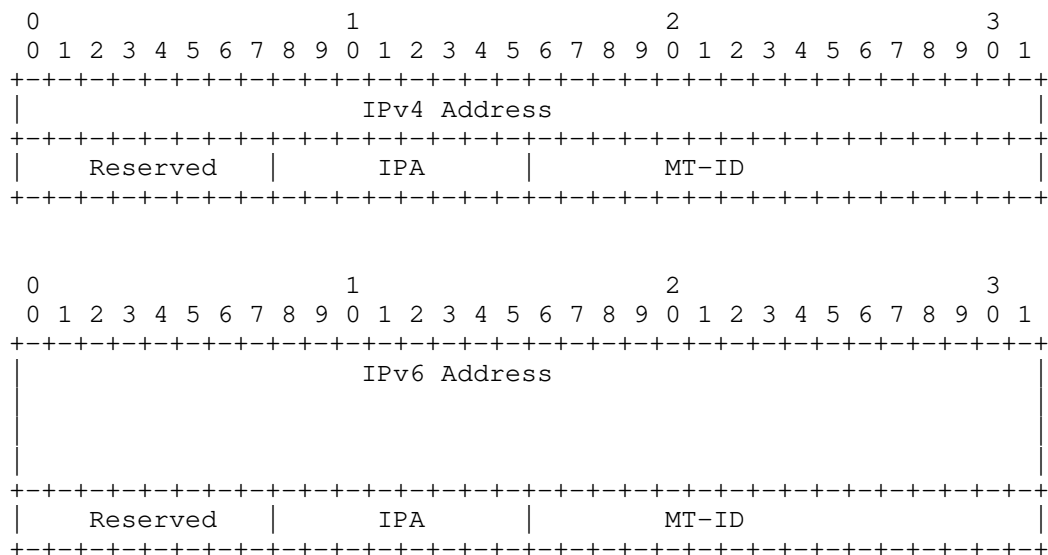


Figure 2: Modified MT IP Address Families Data Format

Where:

**IPv4/IPv6 Address:** An IP address corresponding to "MT IP" and "MT IPv6" address families respectively.

**IPA:** The IGP Algorithm, values are from the IGP Algorithm registry.

**Reserved:** This 8-bit field SHOULD be zero.

#### 4.1.3. MT MP FEC Element

By using extended MT IP Address Family, the resultant MT MP FEC element is to be encoded as follows:

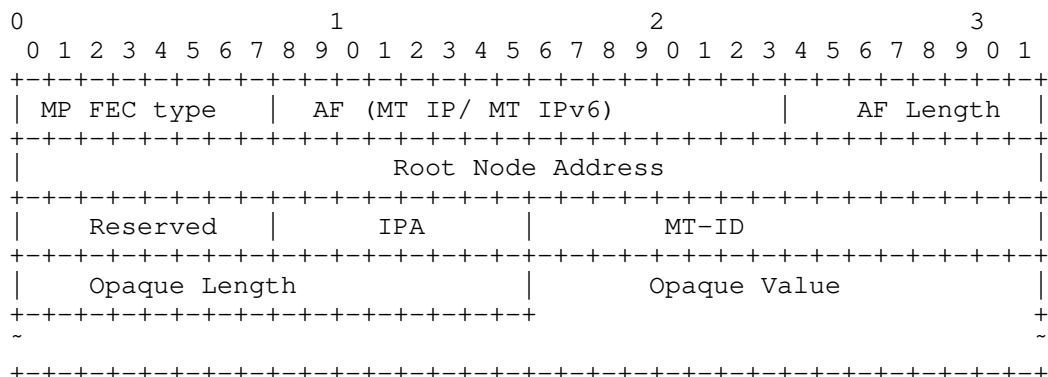


Figure 3: IP MT-Scoped MP FEC Element Format

In the context of this document, the applicable LDP FECs for MT mLDP include:

\* MP FEC Elements:

- P2MP (type 0x6)
- MP2MP-up (type 0x7)
- MP2MP-down (type 0x8)

\* Typed Wildcard FEC Element (type 0x5)

In case of "Typed Wildcard FEC Element", the sub FEC Element type MUST be one of the MP FECs listed above.

This specification allows the use of Topology-scoped mLDP FECs in LDP label and notification messages, as applicable.

#### 4.2. Topology IDs

This document assumes the same definitions and procedures associated with MPLS MT-ID as defined in [RFC7307] specification.

## 5. MT Multipoint Capability

"MT Multipoint Capability" is a new LDP capability, defined in accordance with LDP Capability definition guidelines [RFC5561], that is to be advertised to its peers by an mLDP speaker to announce its capability to support MTR and the procedures specified in this document. This capability MAY be sent either in an Initialization message at the session establishment time, or in a Capability message dynamically during the lifetime of a session (only if "Dynamic Announcement" capability [RFC5561] has been successfully negotiated with the peer).

The format of this capability is as follows:

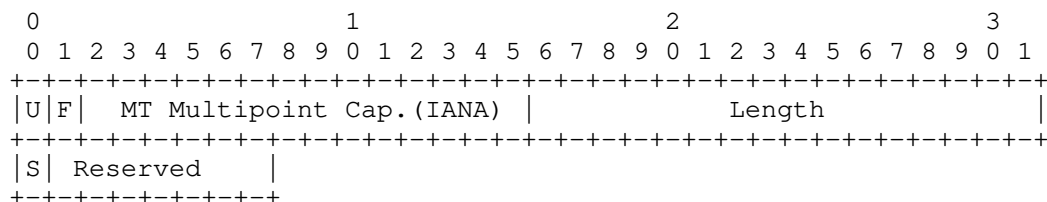


Figure 4: MT Multipoint Capability TLV Format

Where:

U- and F-bits: MUST be 1 and 0, respectively, as per Section 3 of LDP Capabilities [RFC5561].

MT Multipoint Capaillity: TLV type (IANA assigned).

Length: The length (in octets) of TLV. The value of this field MUST be 1 as there is no Capability-specific data [RFC5561] that follows in the TLV.

S-bit: Set to 1 to announce and 0 to withdraw the capability (as per [RFC5561]).

An mLDP speaker that has successfully advertised and negotiated "MT Multipoint" capability MUST support the following:

1. Topology-scoped mLDP FECs in LDP messages (Section 4.1)
2. Topology-scoped mLDP forwarding setup (Section 7)

## 6. MT Applicability on FEC-based features

### 6.1. Typed Wildcard MP FEC Elements

[RFC5918] extends base LDP and defines Typed Wildcard FEC Element framework. Typed Wildcard FEC element can be used in any LDP message to specify a wildcard operation for a given type of FEC.

The MT extensions proposed in document do not require any extension in procedures for Typed Wildcard FEC Element support [RFC5918], and these procedures apply as-is to Multipoint MT FEC wildcarding. Like Typed Wildcard MT Prefix FEC Element, as defined in [RFC7307], the MT extensions allow use of "MT IP" or "MT IPv6" in the Address Family field of the Typed Wildcard MP FEC element in order to use wildcard operations for MP FECs in the context of a given (sub)-topology as identified by the MT-ID and IPA field.

This document proposes following format and encoding for a Typed Wildcard MP FEC element:

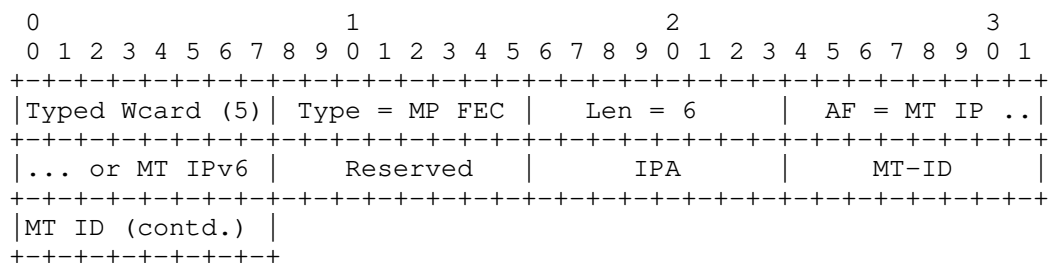


Figure 5: Typed Wildcard MT MP FEC Element

Where:

Type: One of MP FEC Element type (P2MP, MP2MPup, MP2MP-down).

MT ID: MPLS MT ID

IPA: The IGP Algorithm, values are from the IGP Algorithm registry.

The proposed format allows an LSR to perform wildcard MP FEC operations under the scope of a (sub-)topology.

## 6.2. End-of-LIB

[RFC5919] specifies extensions and procedures that allows an LDP speaker to signal its End-of-LIB (i.e. convergence) for a given FEC type towards a peer. MT extensions for MP FEC do not require any change in these procedures and they apply as-is to MT MP FEC elements. This means that an MT mLDP speaker MAY signal its convergence per (sub-)topology using MT Typed Wildcard MP FEC element.

## 7. Topology-Scoped Signaling and Forwarding

Since the {MT-ID, IPA} tuple is part of an mLDP FEC, there is no need to support the concept of multiple (sub-)topology forwarding tables in mLDP. Each MP LSP will be unique due to the tuple being part of the FEC. There is also no need to have specific label forwarding tables per topology, and each MP LSP will have its own unique local label in the table. However, In order to implement MTR in an mLDP network, the selection procedures for upstream LSR and downstream forwarding interface need to be changed.

### 7.1. Upstream LSR selection

The procedures as described in RFC-6388 section-2.4.1.1 depend on the best path to reach the root. When the {MT-ID, IPA} tuple is signaled as part of the FEC, this tuple is used to select the (sub-)topology that must be used to find the best path to the root address. Using the next-hop from this best path, a LDP peer is selected following the procedures as defined in [RFC6388].

### 7.2. Downstream forwarding interface selection

The procedures as described in RFC-6388 section-2.4.1.2 describe how a downstream forwarding interface is selected. In these procedures, any interface leading to the downstream LDP neighbor can be considered as candidate forwarding interface. When the {MT-ID, IPA} tuple is part of the FEC, this is no longer true. An interface must only be selected if it is part of the same (sub-)topology that was signaled in the mLDP FEC element. Besides this restriction, the other procedures in [RFC6388] apply.

## 8. LSP Ping Extensions

[RFC6425] defines procedures to detect data plane failures in Multipoint MPLS LSPs. Section 3.1.2 of [RFC6425] defines new Sub-Types and Sub-TLVs for Multipoint LDP FECs to be sent in "Target FEC Stack" TLV of an MPLS echo request message [RFC4379].

To support LSP ping for MT Multipoint LSPs, this document uses existing sub-types "P2MP LDP FEC Stack" and "MP2MP LDP FEC Stack" defined in [RFC6425]. The proposed extension is to specify "MT IP" or "MT IPv6" in the "Address Family" field, set the "Address Length" field to 8 (for MT IP) or 20 (for MT IPv6), and encode the sub-TLV with additional {MT-ID, IPA} information as an extension to the "Root LSR Address" field. The resultant format of sub-tlv is as follows:

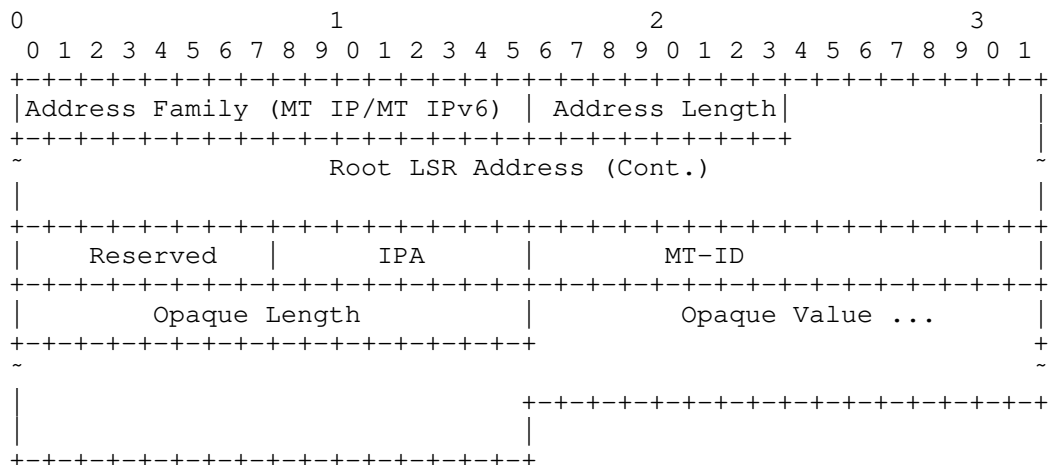


Figure 6: Multipoint LDP FEC Stack Sub-TLV Format for MT

The rules and procedures of using this new sub-TLV in an MPLS echo request message are same as defined for P2MP/MP2MP LDP FEC Stack Sub-TLV in [RFC6425] with only difference being that Root LSR address is now (sub-)topology scoped.

## 9. Security Considerations

This extension to mLDP does not introduce any new security considerations beyond that already apply to the base LDP specification [RFC5036], base mLDP specification [RFC6388], and MPLS security framework [RFC5920].

## 10. IANA Considerations

This document defines a new LDP capability parameter TLV. IANA is requested to assign the lowest available value after 0x0500 from "TLV Type Name Space" in the "Label Distribution Protocol (LDP) Parameters" registry within "Label Distribution Protocol (LDP) Name Spaces" as the new code point for the LDP TLV code point.



Value	Description	Reference	Notes/Registration Date
TBA	MT Multipoint Capability	This document	

Figure 7: IANA Code Point

## 11. Acknowledgments

The authors would like to acknowledge Eric Rosen for his input on this specification.

## 12. References

### 12.1. Normative References

- [I-D.ietf-lsr-flex-algo]  
 Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", Work in Progress, Internet-Draft, draft-ietf-lsr-flex-algo-18, 25 October 2021, <<https://www.ietf.org/archive/id/draft-ietf-lsr-flex-algo-18.txt>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, DOI 10.17487/RFC4379, February 2006, <<https://www.rfc-editor.org/info/rfc4379>>.
- [RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <<https://www.rfc-editor.org/info/rfc4915>>.
- [RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-IS)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <<https://www.rfc-editor.org/info/rfc5120>>.

- [RFC6388] Wijnands, IJ., Ed., Minei, I., Ed., Kompella, K., and B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6388, DOI 10.17487/RFC6388, November 2011, <<https://www.rfc-editor.org/info/rfc6388>>.
- [RFC6425] Saxena, S., Ed., Swallow, G., Ali, Z., Farrel, A., Yasukawa, S., and T. Nadeau, "Detecting Data-Plane Failures in Point-to-Multipoint MPLS - Extensions to LSP Ping", RFC 6425, DOI 10.17487/RFC6425, November 2011, <<https://www.rfc-editor.org/info/rfc6425>>.
- [RFC7307] Zhao, Q., Raza, K., Zhou, C., Fang, L., Li, L., and D. King, "LDP Extensions for Multi-Topology", RFC 7307, DOI 10.17487/RFC7307, July 2014, <<https://www.rfc-editor.org/info/rfc7307>>.

## 12.2. Informative References

- [RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <<https://www.rfc-editor.org/info/rfc5036>>.
- [RFC5561] Thomas, B., Raza, K., Aggarwal, S., Aggarwal, R., and JL. Le Roux, "LDP Capabilities", RFC 5561, DOI 10.17487/RFC5561, July 2009, <<https://www.rfc-editor.org/info/rfc5561>>.
- [RFC5918] Asati, R., Minei, I., and B. Thomas, "Label Distribution Protocol (LDP) 'Typed Wildcard' Forward Equivalence Class (FEC)", RFC 5918, DOI 10.17487/RFC5918, August 2010, <<https://www.rfc-editor.org/info/rfc5918>>.
- [RFC5919] Asati, R., Mohapatra, P., Chen, E., and B. Thomas, "Signaling LDP Label Advertisement Completion", RFC 5919, DOI 10.17487/RFC5919, August 2010, <<https://www.rfc-editor.org/info/rfc5919>>.
- [RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010, <<https://www.rfc-editor.org/info/rfc5920>>.

## Authors' Addresses

IJsbrand Wijnands  
Individual  
Email: [ice@braindump.be](mailto:ice@braindump.be)

Kamran Raza  
Cisco Systems, Inc.  
2000 Innovation Drive  
Kanata ON K2K-3E8  
Canada  
Email: skraza@cisco.com

Mankamana Mishra  
Cisco Systems, Inc.  
821 Alder Drive  
Milpitas, CA 95035  
United States of America  
Email: mankamis@cisco.com

Anuj Budhiraja  
Cisco Systems, Inc.  
821 Alder Drive  
Milpitas, CA 95035  
United States of America  
Email: abudhira@cisco.com

Zhaohui Zhang  
Juniper Networks  
10 Technology Park Dr.  
Westford, MA 01886  
United States of America  
Email: zzhang@juniper.net

Arkadiy Gulko  
Edward Jones wealth management  
United States of America  
Email: Arkadiy.gulko@edwardjones.com