MPLS Transport Profile Data Plane Architecture
draft-ietf-mpls-tp-data-plane-01

Abstract

The Multiprotocol Label Switching (MPLS) Transport Profile (MPLS-TP) is the set of MPLS protocol functions applicable to the construction and operation of packet-switched transport networks. This document specifies the subset of these functions that comprises the MPLS-TP data plane: the architectural layer concerned with the encapsulation and forwarding of packets within an MPLS-TP network.

This document is a product of a joint Internet Engineering Task Force (IETF) / International Telecommunication Union Telecommunication Standardization Sector (ITU-T) effort to include an MPLS Transport Profile within the IETF MPLS and PWE3 architectures to support the capabilities and functionalities of a packet transport network.

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 to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on September 13, 2010.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

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

1. Introduction .................................................. 4  
   1.1. Scope ..................................................... 4  
   1.2. Terminology .............................................. 4  
2. MPLS-TP Packet Encapsulation and Forwarding ................. 5  
3. MPLS-TP Transport Entities ................................. 5  
   3.1. Label Switched Paths ..................................... 6  
      3.1.1. LSP Packet Encapsulation and Forwarding .......... 6  
      3.1.2. LSP Payloads ....................................... 6  
      3.1.3. LSP Types ......................................... 7  
   3.2. Sections ................................................ 8  
   3.3. Pseudowires ............................................. 9  
4. MPLS-TP Generic Associated Channel ....................... 9  
5. Server Layer Considerations ................................ 10  
   5.1. Ethernet Media ......................................... 10  
      5.1.1. Point-to-Point Links .............................. 10  
      5.1.2. Multipoint Links .................................. 11  
6. Security Considerations ................................... 11  
7. IANA Considerations ......................................... 12  
8. References .................................................. 12  
   8.1. Normative References .................................. 12  
   8.2. Informative References ................................ 13  
Authors’ Addresses ............................................. 13
1. Introduction

The MPLS Transport Profile (MPLS-TP) [I-D.ietf-mpls-tp-framework] is the set of protocol functions that meet the requirements [RFC5654] for the application of MPLS to the construction and operation of packet-switched transport networks. Packet transport networks are defined and described in [I-D.ietf-mpls-tp-framework].

This document specifies the subset of protocol functions that comprises the MPLS-TP data plane: the architectural layer concerned with the encapsulation and forwarding of packets within an MPLS-TP network.

This document is a product of a joint Internet Engineering Task Force (IETF) / International Telecommunication Union Telecommunication Standardization Sector (ITU-T) effort to include an MPLS Transport Profile within the IETF MPLS and PWE3 architectures to support the capabilities and functionalities of a packet transport network.

1.1. Scope

This document has the following purposes:

- To identify the data-plane functions within the MPLS Transport Profile;
- To indicate which of these data-plane functions an MPLS-TP implementation is required to support.

Note that the MPLS-TP functions discussed in this document are considered OPTIONAL unless stated otherwise.

1.2. Terminology

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-ACh</td>
<td>Generic Associated Channel</td>
</tr>
<tr>
<td>GAL</td>
<td>G-ACh Label</td>
</tr>
<tr>
<td>LSP</td>
<td>Label Switched Path</td>
</tr>
<tr>
<td>LSR</td>
<td>Label Switching Router</td>
</tr>
<tr>
<td>MAC</td>
<td>Media Access Control</td>
</tr>
<tr>
<td>MPLS-TP</td>
<td>MPLS Transport Profile</td>
</tr>
<tr>
<td>OAM</td>
<td>Operations, Administration and Maintenance</td>
</tr>
<tr>
<td>PW</td>
<td>Pseudowire</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>TTL</td>
<td>Time To Live</td>
</tr>
</tbody>
</table>

Additional definitions and terminology can be found in...
This document defines the encapsulation and forwarding functions applicable to packets traversing an MPLS-TP Label Switched Path (LSP), Pseudowire (PW), or Section (see Section 3 for the definitions of these transport entities). Encapsulation and forwarding functions for packets outside an MPLS-TP LSP, PW, or Section, and mechanisms for delivering packets to or from MPLS-TP LSPs, PWs, and Sections, are outside the scope of this document.

MPLS-TP packet encapsulation and forwarding operates according to the MPLS data-plane architecture described in [RFC3031] and [RFC3032], and the data-plane architectures for Single-Segment Pseudowires [RFC3985], Multi-Segment Pseudowires [RFC5659], and Point-to-Multipoint Pseudowires [I-D.ietf-pwe3-p2mp-pw-requirements], except as noted otherwise in this document.

MPLS-TP forwarding is based on the label that identifies an LSP or PW. The label value specifies the processing operation to be performed by the next hop at that level of encapsulation. A swap of this label is an atomic operation in which the contents of the packet after the swapped label are opaque to the forwarder. The only event that interrupts a swap operation is Time To Live (TTL) expiry.

Further processing to determine the context of a packet occurs when a swap operation is interrupted in this manner by TTL expiry, in which case the BoS label is examined to determine if it is a reserved label.

When a pop operation exposes a specific reserved label, or when the packet is received with the Generic Associated Channel Label (GAL) at the top of the stack. The processing of the Generic Associated Channel Label (GAL) is described in Section 4. Otherwise the packet is forwarded according to the procedures in [RFC3032].

3.  MPLS-TP Transport Entities

The MPLS Transport Profile includes the following data-plane transport entities:

- Label Switched Paths (LSPs)
- Sections
- Pseudowires (PWs)
3.1. Label Switched Paths

MPLS-TP LSPs are ordinary MPLS LSPs as defined in [RFC3031] except as specifically noted otherwise in this document.

3.1.1. LSP Packet Encapsulation and Forwarding

Encapsulation and forwarding of packets traversing MPLS-TP LSPs MUST follow standard MPLS packet encapsulation and forwarding as defined in [RFC3031] and [RFC3032], except as explicitly stated otherwise in this document.

Data-plane support for Internet Protocol (IP) packet encapsulation, addressing, and forwarding is OPTIONAL.

Data-plane Quality of Service capabilities are included in the MPLS-TP in the form of the MPLS Differentiated Services (DiffServ) architecture [RFC3270]. Both E-LSP and L-LSP MPLS DiffServ modes are included. The Traffic Class field (formerly the EXP field) of an MPLS label follows the definition of [RFC5462] and [RFC3270] and MUST be processed according to the rules specified in those documents.

The Pipe and Short Pipe DiffServ tunneling and TTL processing models described in [RFC3270] and [RFC3443] are included in the MPLS-TP. The Uniform model is outside the scope of the MPLS-TP.

Per-platform, per-interface or other context-specific label space [RFC5331] MAY be used for MPLS-TP LSPs. Downstream [RFC3031] or upstream [RFC5331] label allocation schemes MAY be used for MPLS-TP LSPs. Note that the requirements of a particular LSP type may dictate which label spaces or allocation schemes it can use.

Per-packet Equal-Cost Multi-Path (ECMP) load-balancing is outside the scope of the MPLS-TP.

Penultimate Hop Popping (PHP) MUST be disabled by default on MPLS-TP LSPs.

Label merge?

3.1.2. LSP Payloads

The MPLS-TP includes support for the following LSP payload types:

- Network-layer protocol packets
- Pseudowire packets
- LSP

The rules for processing LSP payloads that are network-layer protocol packets SHALL be as specified in [RFC3032].
The rules for processing LSP payloads that are pseudowire packets SHALL be as specified in [RFC3985] and the attendant standards defined by the IETF Pseudowire Emulation Edge-to-Edge (PWE3) Working Group.

Note that the payload of an MPLS-TP LSP may be a packet type that itself contains one or more MPLS labels. This is true, for instance, when the payload is a pseudowire or another MPLS-TP LSP. From the data-plane perspective, however, an MPLS-TP packet is an MPLS packet as specified in [RFC3032], and so in particular has precisely one label stack, and one label in the stack with its S (Bottom of Stack) bit set to 1.

3.1.3. LSP Types

The MPLS-TP includes the following LSP types:

- Point-to-point unidirectional
- Point-to-point associated bidirectional
- Point-to-point co-routed bidirectional
- Point-to-multipoint unidirectional

Point-to-point unidirectional LSPs are supported by the basic MPLS architecture [RFC3031] and are REQUIRED to function in the same manner in the MPLS-TP data plane except as explicitly stated otherwise in this document.

A point-to-point associated bidirectional LSP between LSRs A and B consists of two unidirectional point-to-point LSPs, one from A to B and the other from B to A, which are regarded as a pair providing a single logical bidirectional transport path. The nodes A and B are REQUIRED to be aware of this pairing relationship, but other nodes need not be.

A point-to-point co-routed bidirectional LSP is a point-to-point associated bidirectional LSP with the additional constraint that its two unidirectional component LSPs follow the same path in the network. This means that if one of the component LSPs follows the path through the nodes N0, ..., Nk, originating on N0 and terminating on Nk, then the path of the other component LSP is Nk, ..., N0, and that at each node an ingress interface of one component LSP is an egress interface of the other. In addition, each node along the path is REQUIRED to be aware of the pairing relationship between the component LSPs.

Comment [M15]: The server layer should also be co-routed bidirectional.
3.2. Sections

Two MPLS-TP LSRs are considered to be topologically adjacent at a particular layer n >= 0 of the MPLS-TP LSP hierarchy if there exists a link between them at the next lowest network layer. Such a link, if it exists, will be either an MPLS-TP LSP (if n > 0) or a data-link provided by the underlying server layer network (if n = 0), and is referred to as an MPLS-TP Section at layer n of the MPLS-TP LSP hierarchy. Thus, the links traversed by a layer n+1 MPLS-TP LSP are layer n MPLS-TP sections. Such an LSP is referred to as a client of the section layer, and the section layer as the server layer with respect to its clients.

Note that the MPLS label stack associated with an MPLS-TP section at layer n consists of n labels, in the absence of stack optimisation mechanisms such as PHP. Note also that in order for two LSRs to exchange MPLS-TP control packets over a section, an additional label, the G-ACh Label (GAL) (see Section 4) must appear at the bottom of the label stack.

An MPLS-TP section may provide one or more of the following types of service to its client layer:

- Point-to-point bidirectional
- Point-to-point unidirectional
- Point-to-multipoint unidirectional

The manner in which a section provides such a service is outside the scope of the MPLS-TP.

Note that an LSP of any of the types listed in Section 3.1.3 may serve as a section for a client-layer transport entity as long as it supports the type of service the client requires.
3.3. Pseudowires

The data-plane architectures for Single-Segment Pseudowires [RFC3985], Multi-Segment Pseudowires [RFC5659], and Point-to-Multipoint Pseudowires [I-D.ietf-pwe3-p2mp-pw-requirements], and the associated data-plane pseudowire protocol functions, as defined by the IETF Pseudowire Emulation Edge-to-Edge (PWE3) Working Group, are included in the MPLS-TP. This document specifies no modifications or extensions to pseudowire data-plane architectures or protocols.

4. MPLS-TP Generic Associated Channel

The MPLS Generic Associated Channel (G-ACh) mechanism is specified in [RFC5586] and included in the MPLS-TP. The G-ACh provides an auxiliary logical data channel associated with MPLS-TP Sections, LSPs, and PWs in the data plane. The primary purpose of the G-ACh in the context of MPLS-TP is to support control, management, and OAM traffic associated with MPLS-TP transport entities. The G-ACh MUST NOT be used to transport client layer network traffic in MPLS-TP networks.

For pseudowires, the G-ACh uses the first four bits of the PW control word to provide the initial discrimination between data packets and packets belonging to the associated channel, as described in [RFC4385]. When this first nibble of a packet, immediately following the label at the bottom of stack, has a value of '1', then this packet belongs to a G-ACh. The first 32 bits following the bottom of stack label then have a defined format called an Associated Channel Header (ACH), which further defines the content of the packet. The ACH is therefore both a demultiplexer for G-ACh traffic on the PW, and a discriminator for the type of traffic carried on the G-ACh.

When the control message is carried over a section or an LSP, rather than over a PW, it is necessary to provide an indication in the packet that the payload is something other than a client data packet. This is achieved by including a reserved label with a value of 13 in the label stack. This reserved label is referred to as the G-ACh Label (GAL), and is defined in [RFC5586]. When a GAL is found, it indicates that the payload begins with an ACH. The GAL is thus a demultiplexer for G-ACh traffic on the section or the LSP, and the ACH is a discriminator for the type of traffic carried on the G-ACh. Note however that MPLS-TP forwarding follows the normal MPLS model, and that a GAL is invisible to an LSR unless it is the top label in the label stack. The only other circumstance under which the label stack may be inspected for a GAL is when the TTL has expired.

Comment [M19]:
1) This statement is too vague.
2) PWs may be configured by an OSS.
3) Any references defining data plane behaviour must be normative.
MPLS-TP component that intentionally performs this inspection must assume that it is asynchronous with respect to the forwarding of other packets. All operations on the label stack are in accordance with [RFC3031] and [RFC3032].

5. Server Layer Considerations

This section discusses considerations for support of the MPLS-TP data plane by server layer technologies and media.

In general, the MPLS-TP network has no awareness of the internals of the server layer of which it is a client, requiring only that the server layer be capable of delivering the type of service required by the MPLS-TP transport entities that make use of it. Note that what appears to be a single server layer link to the MPLS-TP network may be a complicated construct underneath, such as an LSP or a collection of underlying links operating as a bundle. Special care may be needed in network design and operation when such constructs are used as a server layer for MPLS-TP.

5.1. Ethernet Media

5.1.1. Point-to-Point Links

When two MPLS-TP nodes are connected by a point-to-point Ethernet link, the question arises as to what destination Ethernet Media Access Control (MAC) address should be specified in Ethernet frames transmitted to the peer node over the link. The problem of determining this address does not arise in IP/MPLS networks because of the presence of the Ethernet Address Resolution Protocol (ARP) [RFC0826] or IP version 6 Neighbor Discovery protocol [RFC4861], which allow the unicast MAC address of the peer device to be learned dynamically.

If existing mechanisms are available in an MPLS-TP network to determine the destination unicast MAC addresses of peer nodes - for example if the network also happens to be an IP/MPLS network - such mechanisms SHOULD be used. The remainder of this section discusses the available options when this is not the case.

One possibility is for each node to be statically configured with the MAC address of its peer. Static MAC address configuration MAY be used in an MPLS-TP network, but can present an administrative burden and lead to operational problems. For example, replacement of an Ethernet interface to resolve a hardware fault when this approach is used requires that the peer node be manually reconfigured with the new MAC address. This is especially problematic if the peer is

Another possibility is to use the Ethernet broadcast address, but this may lead to excessive frame distribution and processing at the Ethernet layer. Broadcast traffic may also be treated specially by some devices and this may not be desirable for MPLS-TP data frames. The preferred approach is therefore to use as the destination MAC address an Ethernet multicast address reserved for MPLS-TP for use over point-to-point links. The address allocated for this purpose by the Internet Assigned Numbers Authority (IANA) is 01-00-5E-XX-XX-XX. An MPLS-TP implementation MUST process Ethernet frames received over a point-to-point link with this destination MAC address by default. Note that this approach is applicable only when the attached Ethernet link is known to be point-to-point. If a link is not known to be point-to-point, the reserved MAC address noted above MUST NOT be used.

Each Ethernet Link Port shall support the following parameters for use by these functions:

a) A Default Link Destination;
b) An adminPointToPointMAC.

The Default Link Destination parameter contains a MAC address to be used in the destination_address parameter of a M_UNITDATA.request at the Ethernet link ISS when a link destination address cannot be derived from the connection identifier parameter of the M_UNITDATA.request from the .... The default value of the Default Link Destination is the Ethernet multicast address reserved for MPLS-TP use over point-to-point links. If the value of the operPointToPointMAC parameter of the Ethernet Link ISS is TRUE then the Default Link Destination parameter is set to the value of the source address parameter of the M_UNITDATA.indication primitive.

The adminPointToPointMAC parameter of the Ethernet link reflects the point-to-point status of the Ethernet link. The default value of the adminPointToPointMAC parameter is ForceFalse. The value may be configured by management to ForceTrue when instantiating an Ethernet link for a point-to-point link. A value of ForceFalse or Auto results in a operPointToPointMAC value of FALSE; a value of ForceTrue results in a operPointToPointMAC value of TRUE. Whenever the operPointToPointMAC parameter transitions to FALSE, the value for the Default Link Destination parameter is set to the 'Ethernet multicast address reserved for MPLS-TP use over point-to-point links'. When the operPointToPointMAC parameter is TRUE, the Default Link Destination parameter is modified by subsequent M_UNITDATA.indications as specified below.

If the value of the operPointToPointMAC parameter of the Ethernet Link ISS is TRUE then the Default Link Destination parameter is set to the value of the source_address parameter of the M_UNITDATA.indication primitive.

The value for the destination address is the contents of the Default Link Destination parameter of the Ethernet Link.

5.1.2. Multipoint Links

Comment [M23]: RFC5332 uses 01-00-5E-XX-XX-XX.

Comment [M24]: The determination of the destination unicast address of a p2p Ethernet link should follow the approach to determine the destination unicast address for a p2p backbone service instance in 802.1ah, clauses 6.10 and 26.4.1. The "adminPointToPointMAC" parameter for the Ethernet link should be set to True in this case. This enforces the use of the source address in the incoming frames from the Ethernet link to be used as destination address in the outgoing frames to the Ethernet link.

Copying some of the text in clause 6.10/802.1ah and adapting this text to fit our case it could read as follows:
When a multipoint Ethernet link serves as a section for a point-to-multipoint MPLS-TP LSP, and multicast destination MAC addressing at the Ethernet layer is used for the LSP, the addressing and encapsulation procedures specified in [RFC5332] SHALL be used.

When a multipoint Ethernet link - that is, a link which is not known to be point-to-point - serves as a section for a point-to-point MPLS-TP LSP, unicast destination MAC addresses must be used for Ethernet frames carrying packets of the LSP. Note that according to the discussion in the previous section, this implies the use of either static MAC address configuration or a protocol that enables peer MAC address discovery.

Also in the case of multipoint links carrying p2p LSPs or PWs it is possible to reuse the mechanism specified in 802.1ah to associate a unicast MAC destination address with each LSP. Where in 802.1ah a relationship is learned between an individual C-MAC address and an individual H-MAC address, in MPLS-TP over multipoint ethernet link a relationship is to be build between a p2p LSP (or PW) and an individual MAC address. This can be done as follows:

a) set up a table with 3 fields for each of the active LSPs: output LSP label value, input LSP label value, individual MAC address
b) fill the input and output LSP label fields for each of the active LSPs
c) read in the input port the values in the source address and LSP label fields of each incoming frame
d) write the value in the source address field of the incoming frame into the individual MAC address field of the row in the table which has the matching input LSP label value.
e) for each outgoing LSP packet, use the output LSP label to lookup the individual MAC address value to be inserted into the destination address of the ethernet frame.
f) if there is no destination address available yet, insert instead the 'Ethernet multicast address reserved for MPLS-TP use over point-to-point links'.

6. Security Considerations

This document serves primarily to specify which aspects of existing MPLS data-plane functionality apply to MPLS-TP. As such it introduces no new security considerations in itself, but the security considerations documented in the specifications to which it refers apply as well to MPLS-TP.
7. IANA Considerations

The authors request that IANA allocate an Ethernet Multicast Address from the Ethernet Multicast Addresses table in the ethernet-numbers registry for use by MPLS-TP LSRs over point-to-point links as described in Section 5.1.1. The entry should specify an address of the form 01-00-5E-XX-XX-XX, a Type Field of 8847/8848, and a usage "MPLS-TP point-to-point (this draft)".

8. References

8.1. Normative References


8.2. Informative References


Authors' Addresses

Dan Frost (editor)
Cisco Systems
Email: danfrost@cisco.com

Stewart Bryant (editor)
Cisco Systems
Email: stbryant@cisco.com

Matthew Bocci (editor)
Alcatel-Lucent
Email: matthew.bocci@alcatel-lucent.com