               MPLS Transport Profile Data Plane Architecture

Abstract

   The Multiprotocol Label Switching 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 Pseudowire Emulation Edge-to-Edge
   (PWE3) architectures to support the capabilities and functionalities
   of a packet transport network.

Status of This Memo

   This is an Internet Standards Track document.

Copyright Notice

Table of Contents

1. Introduction

   The MPLS Transport Profile (MPLS-TP) is the set of functions that
   meet the requirements [RFC5654] for the application of MPLS to the
   construction and operation of packet-switched transport networks.
   MPLS-based packet-switched transport networks, and the overall
   architecture of the MPLS-TP, are defined and described in [RFC5921].
   It is assumed that the reader is familiar with that document.

   This document defines the set of functions that comprise the MPLS-TP
   data plane: the architectural layer concerned with the encapsulation
   and forwarding of packets within an MPLS-TP network. This layer is
   based on the data plane architectures for MPLS ([RFC3031] and
   [RFC3032]) and for pseudowires [RFC3985].

   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:

   o To identify the data plane functions within the MPLS Transport
      Profile; and

   o To indicate which of these data plane functions an MPLS-TP
      implementation is required to support.

   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.

1.2. Terminology

Term	Definition
ACH	Associated Channel Header
G-ACh	Generic Associated Channel
GAL	G-ACh Label
LER	Label Edge Router
LSE	Label Stack Entry
LSP	Label Switched Path
LSR	Label Switching Router
MPLS-TP	MPLS Transport Profile
OAM	Operations, Administration, and Maintenance
PW	Pseudowire
QoS	Quality of Service
S-PE	PW Switching Provider Edge
T-PE	PW Terminating Provider Edge
TTL	Time To Live

Additional definitions and terminology can be found in [RFC5921] and
[RFC5654].

1.3. 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].

2. MPLS-TP Packet Encapsulation and Forwarding

MPLS-TP packet encapsulation and forwarding SHALL operate according
to the MPLS data plane architecture described in [RFC3031] and
[RFC3032] and to the data plane architectures for single-segment
pseudowires and multi-segment pseudowires (see Section 3.3), except
as noted otherwise in this document. The MPLS-TP data plane
satisfies the requirements specified in [RFC5654].

Since an MPLS-TP packet is an MPLS packet as defined in [RFC3031] and
[RFC3032], it will have an associated label stack, and the 'push',
'pop', and 'swap' label processing operations specified in those
documents apply. The label stack represents a hierarchy of Label
Switched Paths (LSPs). A label is pushed to introduce an additional
level of LSP hierarchy and popped to remove it. Such an additional
level may be introduced by any pair of LSRs, whereupon they become
adjacent at this new level, and are then known as Label Edge Routers
(LERs) with respect to the new LSP.

   In contrast to, for example, Section 3.10 of [RFC3031], support for
   Internet Protocol (IP) host and router data plane functionality by
   MPLS-TP interfaces and in MPLS-TP networks is OPTIONAL.

   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 forwarding function. The
   only event that interrupts a swap operation is Time To Live (TTL)
   expiry.

   At an LSR, S-PE, or T-PE, further processing to determine the context
   of a packet occurs when a swap operation is interrupted by TTL
   expiry. If the TTL of an LSP label expires, then the label with the
   S (Bottom of Stack) bit set is inspected to determine if it is a
   reserved label. If it is a reserved label, the packet is processed
   according to the rules of that reserved label. For example, if it is
   a Generic Associated Channel Label (GAL), then it is processed as a
   packet on the Generic Associated Channel (G-ACh); see Section 4. If
   the TTL of a PW expires at an S-PE or T-PE, then the packet is
   examined to determine if a Generic Associated Channel Header (ACH) is
   present immediately below the PW label. If so, then the packet is
   processed as a packet on the G-ACh.

   Similarly, if a pop operation at an LER exposes a reserved label at
   the top of the label stack, then the packet is processed according to
   the rules of that reserved label.

   If no such exception occurs, the packet is forwarded according to the
   procedures in [RFC3031] and [RFC3032].

3. MPLS-TP Transport Entities

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

   o Label Switched Paths (LSPs)

   o sections

   o 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], [RFC3032], [RFC5331], and [RFC5332], except as
   explicitly stated otherwise in this document.

   Data plane Quality of Service capabilities are included in the
   MPLS-TP in the form of Traffic Engineered (TE) LSPs [RFC3209] and 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.

   Except for transient packet reordering that may occur, for example,
   during fault conditions, packets are delivered in order on L-LSPs,
   and on E-LSPs within a specific ordered aggregate.

   The Uniform, Pipe, and Short Pipe Diffserv tunneling and TTL
   processing models described in [RFC3270] and [RFC3443] MAY be used
   for MPLS-TP LSPs. Note, however, that support for the Pipe or Short
   Pipe models is REQUIRED for typical transport applications in which
   the topology and QoS characteristics of the MPLS-TP server layer are
   independent of the client layer. Specific applications MAY place
   further requirements on the Diffserv tunneling and TTL processing
   models an LSP can use.

   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. The requirements of a particular LSP type may, however,
   dictate which label spaces or allocation schemes LSPs of that type
   can use.

   Equal-Cost Multi-Path (ECMP) load-balancing MUST NOT be performed on
   an MPLS-TP LSP. MPLS-TP LSPs as defined in this document MAY operate
   over a server layer that supports load-balancing, but this load-
   balancing MUST operate in such a manner that it is transparent to
   MPLS-TP. This does not preclude the future definition of new MPLS-TP
   LSP types that have different requirements regarding the use of ECMP
   in the server layer.

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

3.1.2. LSP Payloads

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

   o Network-layer protocol packets (including MPLS-labeled packets)

   o Pseudowire packets

   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 defined in the data plane pseudowire specifications (see
   Section 3.3).

   The payload of an MPLS-TP LSP may be a packet that itself contains an
   MPLS label stack. This is true, for instance, when the payload is a
   pseudowire or an MPLS LSP. In such cases, the label stack is
   contiguous between the MPLS-TP LSP and its payload, and exactly one
   LSE in this stack SHALL have the S (Bottom of Stack) bit set to 1.
   This behavior reflects best current practice in MPLS but differs
   slightly from [RFC3032], which uses the S bit to identify when MPLS
   label processing stops and network-layer processing starts.

3.1.3. LSP Types

   The MPLS-TP includes the following LSP types:

   o Point-to-point unidirectional

   o Point-to-point associated bidirectional

   o Point-to-point co-routed bidirectional

   o 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.

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 in each direction follow the same
path (in terms of both nodes and links). An important property of
co-routed bidirectional LSPs is that their unidirectional component
LSPs share fate.

A point-to-multipoint unidirectional LSP functions in the same manner
in the data plane, with respect to basic label processing and packet-
switching operations, as a point-to-point unidirectional LSP, with
one difference: an LSR may have more than one (egress interface,
outgoing label) pair associated with the LSP, and any packet it
transmits on the LSP is transmitted out all associated egress
interfaces. Point-to-multipoint LSPs are described in [RFC4875] and
[RFC5332]. TTL processing and exception handling for point-to-
multipoint LSPs is the same as for point-to-point LSPs and is
described in Section 2.

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
connectivity between them at the next lowest network layer, and if
there is no MPLS layer processing at layer n between the two LSRs
(other than at the LSRs themselves). Such connectivity, 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 is referred to as the server
layer with respect to its clients.

The MPLS label stack associated with an MPLS-TP section at layer n
consists of n labels, in the absence of stack optimization
mechanisms. In order for two LSRs to exchange non-IP 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:

o Point-to-point bidirectional

o Point-to-point unidirectional

o Point-to-multipoint unidirectional

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

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.

A section MUST provide a means of identifying the type of payload it
carries. If the section is a data-link, link-specific mechanisms
such as a protocol type indication in the data-link header MAY be
used. If the section is an LSP, this information MAY be implied by
the LSP label or, if the LSP payload is MPLS-labeled, by the setting
of the S bit. Additional labels MAY also be used if necessary to
distinguish different payload types; see [RFC5921] for examples and
further discussion.

3.3. Pseudowires

The data plane architectures for single-segment pseudowires [RFC3985]
and multi-segment pseudowires [RFC5659] are included in the MPLS-TP.

Data plane processing procedures for pseudowires are defined and
described in a number of IETF documents. Some example pseudowire
data plane procedures include:

o Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over
   an MPLS PSN [RFC4385]

o Encapsulation Methods for Transport of Ethernet over MPLS Networks
   [RFC4448]

o Structure-Agnostic Time Division Multiplexing (TDM) over Packet
   (SAToP) [RFC4553]

o Encapsulation Methods for Transport of PPP/High-Level Data Link
   Control (HDLC) over MPLS Networks [RFC4618]

o Encapsulation Methods for Transport of Frame Relay over
   Multiprotocol Label Switching (MPLS) Networks [RFC4619]

o Encapsulation Methods for Transport of Asynchronous Transfer Mode
   (ATM) over MPLS Networks [RFC4717]

o Pseudowire Emulation Edge-to-Edge (PWE3) Asynchronous Transfer
   Mode (ATM) Transparent Cell Transport Service [RFC4816]

o Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/
   SDH) Circuit Emulation over Packet (CEP) [RFC4842]

    o Structure-Aware Time Division Multiplexed (TDM) Circuit Emulation
       Service over Packet Switched Network (CESoPSN) [RFC5086]

    o Time Division Multiplexing over IP (TDMoIP) [RFC5087]

    o Encapsulation Methods for Transport of Fibre Channel frames Over
       MPLS Networks [FC-ENCAP]

    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
    Operations, Administration, and Maintenance (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 G-ACh traffic.

    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 at
    the bottom of 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. MPLS-TP forwarding follows the normal MPLS model, and thus 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. Normal packet

forwarding MAY continue concurrently with this inspection. All
operations on the label stack are in accordance with [RFC3031] and
[RFC3032].

An application processing a packet received over the G-ACh may
require packet-specific context (such as the receiving interface or
received label stack). Data plane implementations MUST therefore
provide adequate context to the application that is to process a
G-ACh packet. The definition of the context required MUST be
provided as part of the specification of the application using the
G-ACh.

5. Server-Layer Considerations

The MPLS-TP network has no awareness of the internals of the server
layer of which it is a client; it requires 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.

Encapsulation of MPLS-TP packets for transport over specific server-
layer media is outside the scope of this document.

6. Security Considerations

The MPLS data plane (and therefore the MPLS-TP data plane) does not
provide any security mechanisms in and of itself. Client layers that
wish to secure data carried over MPLS-TP transport entities are
REQUIRED to apply their own security mechanisms.

Where management or control plane protocols are used to install
label-switching operations necessary to establish MPLS-TP transport
paths, those protocols are equipped with security features that
network operators may use to securely create the transport paths.

Where enhanced security is desirable, and a trust relationship exists
between an LSR and its peer, the LSR MAY choose to implement the
following policy for the processing of MPLS packets received from one
or more of its neighbors:

    Upon receipt of an MPLS packet, discard the packet unless one of
    the following two conditions holds:

   1. Any MPLS label in the packet's label stack processed at the
       receiving LSR, such as an LSP or PW label, has a label value
       that the receiving LSR has distributed to that neighbor; or

   2. Any MPLS label in the packet's label stack processed at the
       receiving LSR, such as an LSP or PW label, has a label value
       that the receiving LSR has previously distributed to the peer
       beyond that neighbor (i.e., when it is known that the path
       from the system to which the label was distributed to the
       receiving system is via that neighbor).

   Further details of MPLS and MPLS-TP security can be found in
   [RFC5921] and [RFC5920].

7. References

7.1. Normative References

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

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

   [RFC3032]   Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y.,
               Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack
               Encoding", RFC 3032, January 2001.

   [RFC3209]   Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
               and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
               Tunnels", RFC 3209, December 2001.

   [RFC3270]   Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen,
               P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-
               Protocol Label Switching (MPLS) Support of Differentiated
               Services", RFC 3270, May 2002.

   [RFC3443]   Agarwal, P. and B. Akyol, "Time To Live (TTL) Processing
               in Multi-Protocol Label Switching (MPLS) Networks",
               RFC 3443, January 2003.

   [RFC4385]   Bryant, S., Swallow, G., Martini, L., and D. McPherson,
               "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word
               for Use over an MPLS PSN", RFC 4385, February 2006.

   [RFC4448]   Martini, L., Rosen, E., El-Aawar, N., and G. Heron,
               "Encapsulation Methods for Transport of Ethernet over
               MPLS Networks", RFC 4448, April 2006.

   [RFC4553]   Vainshtein, A. and YJ. Stein, "Structure-Agnostic Time
               Division Multiplexing (TDM) over Packet (SAToP)",
               RFC 4553, June 2006.

   [RFC4618]   Martini, L., Rosen, E., Heron, G., and A. Malis,
               "Encapsulation Methods for Transport of PPP/High-Level
               Data Link Control (HDLC) over MPLS Networks", RFC 4618,
               September 2006.

   [RFC4619]   Martini, L., Kawa, C., and A. Malis, "Encapsulation
               Methods for Transport of Frame Relay over Multiprotocol
               Label Switching (MPLS) Networks", RFC 4619,
               September 2006.

   [RFC4717]   Martini, L., Jayakumar, J., Bocci, M., El-Aawar, N.,
               Brayley, J., and G. Koleyni, "Encapsulation Methods for
               Transport of Asynchronous Transfer Mode (ATM) over MPLS
               Networks", RFC 4717, December 2006.

   [RFC4816]   Malis, A., Martini, L., Brayley, J., and T. Walsh,
               "Pseudowire Emulation Edge-to-Edge (PWE3) Asynchronous
               Transfer Mode (ATM) Transparent Cell Transport Service",
               RFC 4816, February 2007.

   [RFC4842]   Malis, A., Pate, P., Cohen, R., and D. Zelig,
               "Synchronous Optical Network/Synchronous Digital
               Hierarchy (SONET/SDH) Circuit Emulation over Packet
               (CEP)", RFC 4842, April 2007.

   [RFC4875]   Aggarwal, R., Papadimitriou, D., and S. Yasukawa,
               "Extensions to Resource Reservation Protocol - Traffic
               Engineering (RSVP-TE) for Point-to-Multipoint TE Label
               Switched Paths (LSPs)", RFC 4875, May 2007.

   [RFC5331]   Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream
               Label Assignment and Context-Specific Label Space",
               RFC 5331, August 2008.

   [RFC5332]   Eckert, T., Rosen, E., Aggarwal, R., and Y. Rekhter,
               "MPLS Multicast Encapsulations", RFC 5332, August 2008.

   [RFC5462]   Andersson, L. and R. Asati, "Multiprotocol Label
               Switching (MPLS) Label Stack Entry: "EXP" Field Renamed
               to "Traffic Class" Field", RFC 5462, February 2009.

   [RFC5586]   Bocci, M., Vigoureux, M., and S. Bryant, "MPLS Generic
               Associated Channel", RFC 5586, June 2009.

   [RFC5654] Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N.,
              and S. Ueno, "Requirements of an MPLS Transport Profile",
              RFC 5654, September 2009.

7.2. Informative References

   [FC-ENCAP] Black, D. and L. Dunbar, "Encapsulation Methods for
              Transport of Fibre Channel frames Over MPLS Networks",
              Work in Progress, June 2010.

   [RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-
              Edge (PWE3) Architecture", RFC 3985, March 2005.

   [RFC5086] Vainshtein, A., Sasson, I., Metz, E., Frost, T., and P.
              Pate, "Structure-Aware Time Division Multiplexed (TDM)
              Circuit Emulation Service over Packet Switched Network
              (CESoPSN)", RFC 5086, December 2007.

   [RFC5087] Stein, Y(J)., Shashoua, R., Insler, R., and M. Anavi,
              "Time Division Multiplexing over IP (TDMoIP)", RFC 5087,
              December 2007.

   [RFC5659] Bocci, M. and S. Bryant, "An Architecture for Multi-
              Segment Pseudowire Emulation Edge-to-Edge", RFC 5659,
              October 2009.

   [RFC5920] Fang, L., "Security Framework for MPLS and GMPLS
              Networks", RFC 5920, July 2010.

   [RFC5921] Bocci, M., Bryant, S., Frost, D., Levrau, L., and L.
              Berger, "A Framework for MPLS in Transport Networks",
              RFC 5921, July 2010.

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