Internet Engineering Task Force (IETF)
Request for Comments: 6371
Category: Informational
ISSN: 2070-1721

I. Busi, Ed.
Alcatel-Lucent
D. Allan, Ed.
Ericsson

September 2011

Operations, Administration, and Maintenance Framework for MPLS-Based Transport Networks

Abstract

The Transport Profile of Multiprotocol Label Switching (MPLS-TP) is a packet-based transport technology based on the MPLS Traffic Engineering (MPLS-TE) and pseudowire (PW) data-plane architectures.

This document describes a framework to support a comprehensive set of Operations, Administration, and Maintenance (OAM) procedures that fulfill the MPLS-TP OAM requirements for fault, performance, and protection-switching management and that do not rely on the presence of a control plane.

This document is a product of a joint Internet Engineering Task Force (IETF) / International Telecommunications 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 as defined by the ITU-T.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6371.

Table of Contents

1. Introduction

As noted in the MPLS Transport Profile (MPLS-TP) framework RFCs (RFC 5921 [8] and RFC 6215 [9]), MPLS-TP is a packet-based transport technology based on the MPLS Traffic Engineering (MPLS-TE) and pseudowire (PW) data-plane architectures defined in RFC 3031 [1], RFC 3985 [2], and RFC 5659 [4].

MPLS-TP utilizes a comprehensive set of Operations, Administration, and Maintenance (OAM) procedures for fault, performance, and protection-switching management that do not rely on the presence of a control plane.

In line with [15], existing MPLS OAM mechanisms will be used wherever possible, and extensions or new OAM mechanisms will be defined only where existing mechanisms are not sufficient to meet the requirements. Some extensions discussed in this framework may end up

as aspirational capabilities and may be determined to be not tractably realizable in some implementations. Extensions do not deprecate support for existing MPLS OAM capabilities.

The MPLS-TP OAM framework defined in this document provides a protocol-neutral description of the required OAM functions and of the data-plane OAM architecture to support a comprehensive set of OAM procedures that satisfy the MPLS-TP OAM requirements of RFC 5860 [11]. In this regard, it defines similar OAM functionality as for existing Synchronous Optical Network / Synchronous Digital Hierarchy (SONET/SDH) and Optical Transport Network (OTN) OAM mechanisms (e.g., [19]).

The MPLS-TP OAM framework is applicable to Sections, Label Switched Paths (LSPs), Multi-Segment Pseudowires (MS-PWs), and Sub-Path Maintenance Elements (SPMEs). It supports co-routed and associated bidirectional P2P transport paths as well as unidirectional P2P and P2MP transport paths.

OAM packets that instrument a particular direction of a transport path are subject to the same forwarding treatment (i.e., fate-share) as the user data packets and in some cases, where Explicitly TC-encoded-PSC LSPs (E-LSPs) are employed, may be required to have common per-hop behavior (PHB) Scheduling Class (PSC) End-to-End (E2E) with the class of traffic monitored. In case of Label-Only-Inferred-PSC LSP (L-LSP), only one class of traffic needs to be monitored, and therefore the OAM packets have common PSC with the monitored traffic class.

OAM packets can be distinguished from the used data packets using the Generic Associated Channel Label (GAL) and Associated Channel Header (ACH) constructs of RFC 5586 [7] for LSP, SPME, and Section, or the ACH construct of RFC 5085 [3] and RFC 5586 [7] for (MS-)PW. OAM packets are never fragmented and are not combined with user data in the same packet payload.

This framework makes certain assumptions as to the utility and frequency of different classes of measurement that naturally suggest different functions are implemented as distinct OAM flows or packets. This is dictated by the combination of the class of problem being detected and the need for timeliness of network response to the problem. For example, fault detection is expected to operate on an entirely different time base than performance monitoring, which is also expected to operate on an entirely different time base than in-band management transactions.

The remainder of this memo is structured as follows:

Section 2 covers the definitions and terminology used in this memo.

Section 3 describes the functional component that generates and processes OAM packets.

Section 4 describes the reference models for applying OAM functions to Sections, LSP, MS-PW, and their SPMEs.

Sections 5, 6, and 7 provide a protocol-neutral description of the OAM functions, defined in RFC 5860 [11], aimed at clarifying how the OAM protocol solutions will behave to achieve their functional objectives.

Section 8 discusses the security implications of OAM protocol design in the MPLS-TP context.

The OAM protocol solutions designed as a consequence of this document are expected to comply with the functional behavior described in Sections 5, 6, and 7. Alternative solutions to required functional behaviors may also be defined.

OAM specifications following this OAM framework may be provided in different documents to cover distinct OAM functions.

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 as defined by the ITU-T.

2. Conventions Used in This Document

2.1. Terminology

AC Attachment Circuit

AIS Alarm Indication Signal

CC Continuity Check

CC-V Continuity Check and Connectivity Verification

CV Connectivity Verification

DBN Domain Border Node

E-LSP	Explicitly TC-encoded-PSC LSP
ICC	ITU Carrier Code
LER	Label Edge Router
LKR	Lock Report
L-LSP	Label-Only-Inferred-PSC LSP
LM	Loss Measurement
LME	LSP Maintenance Entity
LMEG	LSP ME Group
LSP	Label Switched Path
LSR	Label Switching Router
LSME	LSP SPME ME
LSMEG	LSP SPME ME Group
ME	Maintenance Entity
MEG	Maintenance Entity Group
MEP	Maintenance Entity Group End Point
MIP	Maintenance Entity Group Intermediate Point
NMS	Network Management System
PE	Provider Edge
PHB	Per-Hop Behavior
PM	Performance Monitoring
PME	PW Maintenance Entity
PMEG	PW ME Group
PSC	PHB Scheduling Class
PSME	PW SPME ME

PSMEG PW SPME ME Group

PW Pseudowire

SLA Service Level Agreement

SME Section Maintenance Entity

SMEG Section ME Group

SPME Sub-Path Maintenance Element

S-PE Switching Provider Edge

TC Traffic Class

T-PE Terminating Provider Edge

2.2. Definitions

This document uses the terms defined in RFC 5654 [5].

This document uses the term 'per-hop behavior' as defined in RFC 2474 [16].

This document uses the term 'LSP' to indicate either a service LSP or a transport LSP (as defined in RFC 5921 [8]).

This document uses the term 'Section' exclusively to refer to the n=0 case of the term 'Section' defined in RFC 5960 [10].

This document uses the term 'Sub-Path Maintenance Element (SPME)' as defined in RFC 5921 [8].

This document uses the term 'traffic profile' as defined in RFC 2475 [13].

Where appropriate, the following definitions are aligned with ITU-T recommendation Y.1731 [21] in order to have a common, unambiguous terminology. They do not however intend to imply a certain implementation but rather serve as a framework to describe the necessary OAM functions for MPLS-TP.

Adaptation function: The adaptation function is the interface between the client (sub-)layer and the server (sub-)layer.

Branch Node: A node along a point-to-multipoint transport path that is connected to more than one downstream node.

Bud Node: A node along a point-to-multipoint transport path that is at the same time a branch node and a leaf node for this transport path.

Data-plane loopback: An out-of-service test where a transport path at either an intermediate or terminating node is placed into a data-plane loopback state, such that all traffic (including both payload and OAM) received on the looped back interface is sent on the reverse direction of the transport path.

Note: The only way to send an OAM packet to a node that has been put into data-plane loopback mode is via Time to Live (TTL) expiry, irrespective of whether the node is hosting MIPs or MEPs.

Domain Border Node (DBN): An intermediate node in an MPLS-TP LSP that is at the boundary between two MPLS-TP OAM domains. Such a node may be present on the edge of two domains or may be connected by a link to the DBN at the edge of another OAM domain.

Down MEP: A MEP that receives OAM packets from, and transmits them towards, the direction of a server layer.

Forwarding Engine: An abstract functional component, residing in an LSR, that forwards the packets from an ingress interface toward the egress interface(s).

In-Service: The administrative status of a transport path when it is unlocked.

Interface: An interface is the attachment point to a server (sub-)layer, e.g., a MPLS-TP Section or MPLS-TP tunnel.

Intermediate Node: An intermediate node transits traffic for an LSP or a PW. An intermediate node may originate OAM flows directed to downstream intermediate nodes or MEPs.

Loopback: See data-plane loopback and OAM loopback definitions.

Maintenance Entity (ME): Some portion of a transport path that requires management bounded by two points (called MEPs), and the relationship between those points to which maintenance and monitoring operations apply (details in Section 3.1).

Maintenance Entity Group (MEG): The set of one or more maintenance entities that maintain and monitor a section or a transport path in an OAM domain.

MEP: A MEG End Point (MEP) is capable of initiating (source MEP) and terminating (sink MEP) OAM packets for fault management and performance monitoring. MEPs define the boundaries of an ME (details in Section 3.3).

MIP: A MEG intermediate point (MIP) terminates and processes OAM packets that are sent to this particular MIP and may generate OAM packets in reaction to received OAM packets. It never generates unsolicited OAM packets itself. A MIP resides within a MEG between MEPs (details in Section 3.3).

OAM domain: A domain, as defined in [5], whose entities are grouped for the purpose of keeping the OAM confined within that domain. An OAM domain contains zero or more MEGs.

Note: Within the rest of this document, the term "domain" is used to indicate an "OAM domain".

OAM flow: The set of all OAM packets originating with a specific source MEP that instrument one direction of a MEG (or possibly both in the special case of data-plane loopback).

OAM loopback: The capability of a node to be directed by a received OAM packet to generate a reply back to the sender. OAM loopback can work in-service and can support different OAM functions (e.g., bidirectional on-demand connectivity verification).

OAM Packet: A packet that carries OAM information between MEPs and/or MIPs in a MEG to perform some OAM functionality (e.g., connectivity verification).

Originating MEP: A MEP that originates an OAM transaction packet (toward a target MIP/MEP) and expects a reply, either in-band or out-of-band, from that target MIP/MEP. The originating MEP always generates the OAM request packets in-band and expects and processes only OAM reply packets returned by the target MIP/MEP.

Out-of-Service: The administrative status of a transport path when it is locked. When a path is in a locked condition, it is blocked from carrying client traffic.

Path Segment: It is either a segment or a concatenated segment, as defined in RFC 5654 [5].

Signal Degrade: A condition declared by a MEP when the data forwarding capability associated with a transport path has deteriorated, as determined by performance monitoring (PM). See also ITU-T recommendation G.806 [14].

Signal Fail: A condition declared by a MEP when the data forwarding capability associated with a transport path has failed, e.g., loss of continuity. See also ITU-T recommendation G.806 [14].

Sink MEP: A MEP acts as a sink MEP for an OAM packet when it terminates and processes the packets received from its associated MEG.

Source MEP: A MEP acts as source MEP for an OAM packet when it originates and inserts the packet into the transport path for its associated MEG.

Tandem Connection: A tandem connection is an arbitrary part of a transport path that can be monitored (via OAM) independent of the end-to-end monitoring (OAM). The tandem connection may also include the forwarding engine(s) of the node(s) at the boundaries of the tandem connection. Tandem connections may be nested but cannot overlap. See also ITU-T recommendation G.805 [20].

Target MEP/MIP: A MEP or a MIP that is targeted by OAM transaction packets and that replies to the originating MEP that initiated the OAM transactions. The target MEP or MIP can reply either in-band or out-of-band. The target sink MEP function always receives the OAM request packets in-band, while the target source MEP function only generates the OAM reply packets that are sent in-band.

Up MEP: A MEP that transmits OAM packets towards, and receives them from, the direction of the forwarding engine.

3. Functional Components

MPLS-TP is a packet-based transport technology based on the MPLS and PW data plane architectures ([1], [2], and [4]) and is capable of transporting service traffic where the characteristics of information transfer between the transport path end points can be demonstrated to comply with certain performance and quality guarantees.

In order to describe the required OAM functionality, this document introduces a set of functional components.

3.1. Maintenance Entity and Maintenance Entity Group

MPLS-TP OAM operates in the context of Maintenance Entities (MEs) that define a relationship between two points of a transport path to which maintenance and monitoring operations apply. The two points that define a maintenance entity are called Maintenance Entity Group End Points (MEPs). The collection of one or more MEs that belongs to the same transport path and that are maintained and monitored as a

group are known as a Maintenance Entity Group (MEG). In between MEPs, there are zero or more intermediate points, called Maintenance Entity Group Intermediate Points (MIPs). MEPs and MIPs are associated with the MEG and can be shared by more than one ME in a MEG.

An abstract reference model for an ME is illustrated in Figure 1 below.

Figure 1: ME Abstract Reference Model

The instantiation of this abstract model to different MPLS-TP entities is described in Section 4. In Figure 1, nodes A and D can be Label Edge Routers (LERs) for an LSP or the Terminating Provider Edges (T-PEs) for an MS-PW, nodes B and C are LSRs for an LSP or Switching PEs (S-PEs) for an MS-PW. MEPs reside in nodes A and D, while MIPs reside in nodes B and C and may reside in A and D. The links connecting adjacent nodes can be physical links, (sub-)layer LSPs/SPMEs, or server-layer paths.

This functional model defines the relationships between all OAM entities from a maintenance perspective and it allows each Maintenance Entity to provide monitoring and management for the (sub-)layer network under its responsibility and efficient localization of problems.

An MPLS-TP Maintenance Entity Group may be defined to monitor the transport path for fault and/or performance management.

The MEPs that form a MEG bound the scope of an OAM flow to the MEG (i.e., within the domain of the transport path that is being monitored and managed). There are two exceptions to this:

1) A misbranching fault may cause OAM packets to be delivered to a MEP that is not in the MEG of origin.

2) An out-of-band return path may be used between a MIP or a MEP and the originating MEP.

In case of a unidirectional point-to-point transport path, a single unidirectional Maintenance Entity is defined to monitor it.

In case of associated bidirectional point-to-point transport paths, two independent unidirectional Maintenance Entities are defined to independently monitor each direction. This has implications for transactions that terminate at or query a MIP, as a return path from MIP to the originating MEP does not necessarily exist in the MEG.

In case of co-routed bidirectional point-to-point transport paths, a single bidirectional Maintenance Entity is defined to monitor both directions congruently.

In case of unidirectional point-to-multipoint transport paths, a single unidirectional Maintenance Entity for each leaf is defined to monitor the transport path from the root to that leaf.

In all cases, portions of the transport path may be monitored by the instantiation of SPMEs (see Section 3.2).

The reference model for the P2MP MEG is represented in Figure 2.

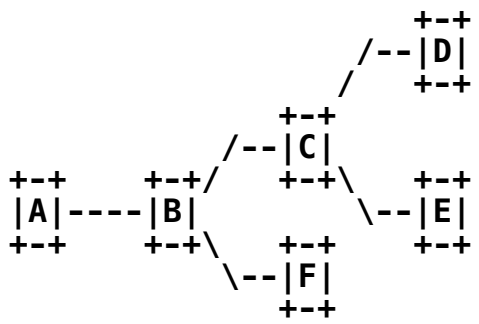


Figure 2: Reference Model for P2MP MEG

In the case of P2MP transport paths, the OAM measurements are independent for each ME (A-D, A-E, and A-F):

- Fault conditions - some faults may impact more than one ME depending on where the failure is located;
- Packet loss - packet dropping may impact more than one ME depending from where the packets are lost;
- Packet delay - will be unique per ME.

Each leaf (i.e., D, E, and F) terminates OAM flows to monitor the ME between itself and the root while the root (i.e., A) generates OAM packets common to all the MEs of the P2MP MEG. All nodes may implement a MIP in the corresponding MEG.

3.2. MEG Nesting: SPMEs and Tandem Connection Monitoring

In order to verify and maintain performance and quality guarantees, there is a need to apply OAM functionality not only on a transport path granularity (e.g., LSP or MS-PW), but also on arbitrary parts of transport paths, defined as tandem connections, between any two arbitrary points along a transport path.

Sub-Path Maintenance Elements (SPMEs), as defined in [8], are hierarchical LSPs instantiated to provide monitoring of a portion of a set of transport paths (LSPs or MS-PWs) that follow the same path between the ingress and the egress of the SPME. The operational aspects of instantiating SPMEs are out of scope of this memo.

SPMEs can also be employed to meet the requirement to provide tandem connection monitoring (TCM), as defined by ITU-T Recommendation G.805 [20].

TCM for a given path segment of a transport path is implemented by creating an SPME that has a 1:1 association with the path segment of the transport path that is to be monitored.

In the TCM case, this means that the SPME used to provide TCM can carry one and only one transport path, thus allowing direct correlation between all fault management and performance monitoring information gathered for the SPME and the monitored path segment of the end-to-end transport path.

There are a number of implications to this approach:

1) The SPME would use the uniform model [23] of Traffic Class (TC) code point copying between sub-layers for Diffserv such that the E2E markings and PHB treatment for the transport path were preserved by the SPMEs.

2) The SPME normally would use the short-pipe model for TTL handling [6] (no TTL copying between sub-layers) such that the TTL distance to the MIPs for the E2E entity would not be impacted by the presence of the SPME, but it should be possible for an operator to specify use of the uniform model.

Note that points 1 and 2 above assume that the TTL copying mode and TC copying modes are independently configurable for an LSP.

The TTL distance to the MIPs plays a critical role for delivering packets to these MIPs as described in Section 3.4.

There are specific issues with the use of the uniform model of TTL copying for an SPME:

1. A MIP in the SPME sub-layer is not part of the transport-path MEG; hence, only an out-of-band return path for OAM originating in the transport-path MEG that addressed an SPME MIP might be available.

2. The instantiation of a lower-level MEG or protection-switching actions within a lower-level MEG may change the TTL distances to MIPs in the higher-level MEGs.

The end points of the SPME are MEPs and limit the scope of an OAM flow within the MEG that the MEPs belong to (i.e., within the domain of the SPME that is being monitored and managed).

When considering SPMEs, it is important to consider that the following properties apply to all MPLS-TP MEGs (regardless of whether they instrument LSPs, SPMEs, or MS-PWs):

- They can be nested but not overlapped, e.g., a MEG may cover a path segment of another MEG and may also include the forwarding engine(s) of the node(s) at the edge(s) of the path segment. However, when MEGs are nested, the MEPs and MIPs in the SPME are no longer part of the encompassing MEG.

- It is possible that MEPs of MEGs that are nested reside on a single node but again are implemented in such a way that they do not overlap.

- Each OAM flow is associated with a single MEG.

- When an SPME is instantiated after the transport path has been instantiated, the TTL distance to the MIPs may change for the short-pipe model of TTL copying, and may change for the uniform model if the SPME is not co-routed with the original path.

3.3. MEG End Points (MEPs)

MEG End Points (MEPs) are the source and sink points of a MEG. In the context of an MPLS-TP LSP, only LERs can implement MEPs, while in the context of an SPME, any LSR of the MPLS-TP LSP can be an LER of SPMEs that contributes to the overall monitoring infrastructure of the transport path. Regarding PWs, only T-PEs can implement MEPs; while for SPMEs supporting one or more PWs, both T-PEs and S-PEs can implement SPME MEPs. Any MPLS-TP LSR can implement a MEP for an MPLS-TP Section.

MEPs are responsible for originating almost all of the proactive and on-demand monitoring OAM functionality for the MEG. There is a separate class of notifications (such as Lock Report (LKR) and Alarm Indication Signal (AIS)) that are originated by intermediate nodes and triggered by server-layer events. A MEP is capable of originating and terminating OAM packets for fault management and performance monitoring. These OAM packets are carried within the Generic Associated Channel (G-ACh) with the proper encapsulation and an appropriate channel type as defined in RFC 5586 [7]. A MEP terminates all the OAM packets it receives from the MEG it belongs to and silently discards those that do not. (Note that in the particular case of Connectivity Verification (CV) processing, a CV packet from an incorrect MEG will result in a mis-connectivity defect and there are further actions taken.) The MEG the OAM packet belongs to is associated with the MPLS or PW label, whether the label is used to infer the MEG or the content of the OAM packet is an implementation choice. In the case of an MPLS-TP Section, the MEG is inferred from the port on which an OAM packet was received with the GAL at the top of the label stack.

OAM packets may require the use of an available "out-of-band" return path (as defined in [8]). In such cases, sufficient information is required in the originating transaction such that the OAM reply packet can be constructed and properly forwarded to the originating MEP (e.g., IP address).

Each OAM solution document will further detail the applicability of the tools it defines as a proactive or on-demand mechanism as well as its usage when:

- The "in-band" return path exists and it is used.
- An "out-of-band" return path exists and it is used.
- Any return path does not exist or is not used.

Once a MEG is configured, the operator can configure which proactive OAM functions to use on the MEG, but the MEPs are always enabled.

MEPs terminate all OAM packets received from the associated MEG. As the MEP corresponds to the termination of the forwarding path for a MEG at the given (sub-)layer, OAM packets never leak outside of a MEG in a properly configured fault-free implementation.

A MEP of an MPLS-TP transport path coincides with transport path termination and monitors it for failures or performance degradation (e.g., based on packet counts) in an end-to-end scope. Note that both the source MEP and sink MEP coincide with transport paths' source and sink terminations.

The MEPs of an SPME are not necessarily coincident with the termination of the MPLS-TP transport path. They are used to monitor a path segment of the transport path for failures or performance degradation (e.g., based on packet counts) only within the boundary of the MEG for the SPME.

An MPLS-TP sink MEP passes a fault indication to its client (sub-)layer network as a consequent action of fault detection. When the client layer is not MPLS-TP, the consequent actions in the client layer (e.g., ignore or generate client-layer-specific OAM notifications) are outside the scope of this document.

A node hosting a MEP can either support per-node MEP or per-interface MEP(s). A per-node MEP resides in an unspecified location within the node, while a per-interface MEP resides on a specific side of the forwarding engine. In particular, a per-interface MEP is called an "Up MEP" or a "Down MEP" depending on its location relative to the forwarding engine. An "Up MEP" transmits OAM packets towards, and receives them from, the direction of the forwarding engine, while a "Down MEP" receives OAM packets from, and transmits them towards, the direction of a server layer.

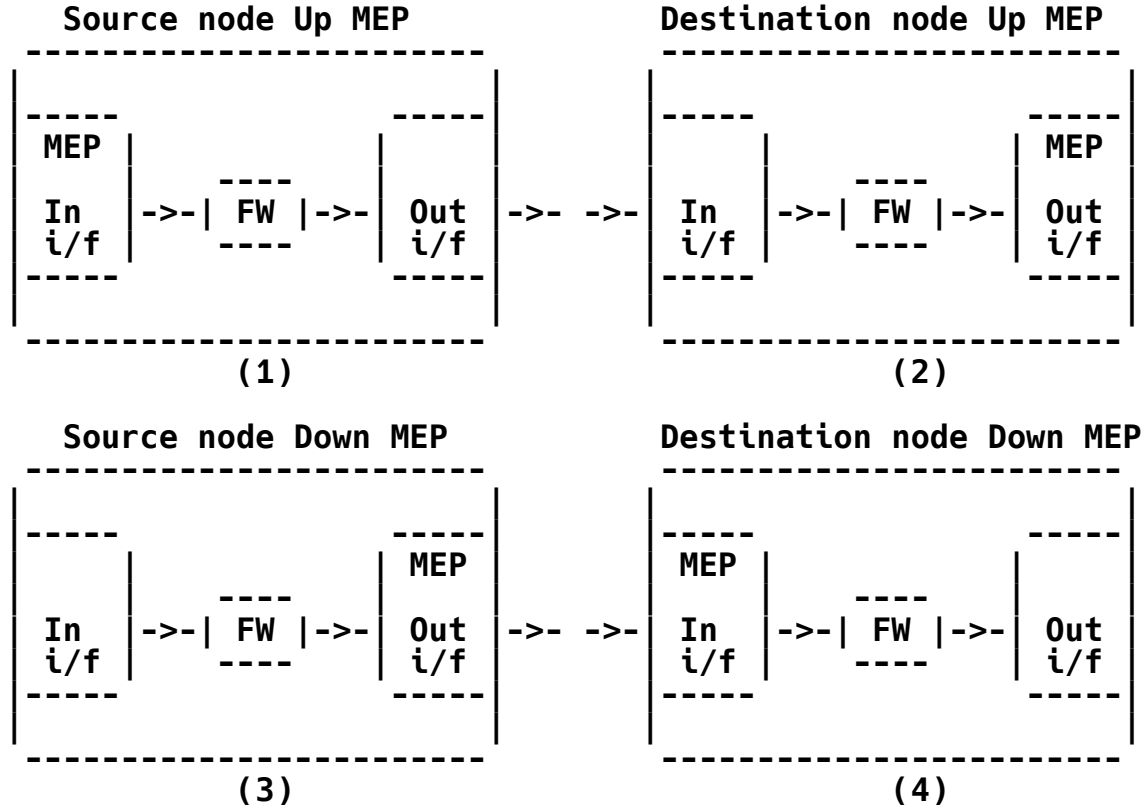


Figure 3: Examples of Per-Interface MEPs

Figure 3 describes four examples of per-interface Up MEPs: an Up Source MEP in a source node (case 1), an Up Sink MEP in a destination node (case 2), a Down Source MEP in a source node (case 3), and a Down Sink MEP in a destination node (case 4).

The usage of per-interface Up MEPs extends the coverage of the ME for both fault and performance monitoring closer to the edge of the domain and determines that the location of a failure or performance degradation is within a node or on a link between two adjacent nodes.

Each OAM solution document will further detail the implications of the tools it defines when used with per-interface or per-node MEPs, if necessary.

It may occur that multiple MEPs for the same MEG are on the same node, and are all Up MEPs, each on one side of the forwarding engine, such that the MEG is entirely internal to the node.

It should be noted that an ME may span nodes that implement per-node MEPs and per-interface MEPs. This guarantees backward compatibility with most of the existing LSRs that can implement only a per-node MEP. In fact, in many current implementations, label operations are largely performed on the ingress interface; hence, the exposure of the GAL as top label will occur at the ingress interface.

Note that a MEP can only exist at the beginning and end of a (sub-)layer in MPLS-TP. If there is a need to monitor some portion of that LSP or PW, a new sub-layer (in the form of an SPME) must be created that permits MEPs and associated MEGs to be created.

In the case where an intermediate node sends an OAM packet to a MEP, it uses the top label of the stack at that point.

3.4. MEG Intermediate Points (MIPs)

A MEG Intermediate Point (MIP) is a function located at a point between the MEPs of a MEG for a PW, LSP, or SPME.

A MIP is capable of reacting to some OAM packets and forwarding all the other OAM packets while ensuring fate-sharing with user data packets. However, a MIP does not initiate unsolicited OAM packets, but may be addressed by OAM packets initiated by one of the MEPs of the MEG. A MIP can generate OAM packets only in response to OAM packets that it receives from the MEG it belongs to. The OAM packets generated by the MIP are sent to the originating MEP.

An intermediate node within a MEG can either:

- support per-node MIPs (i.e., a single MIP per node in an unspecified location within the node); or

- support per-interface MIPs (i.e., two or more MIPs per node on both sides of the forwarding engine).

Support of per-interface or per-node MIPs is an implementation choice. It is also possible that a node could support per-interface MIPs on some MEGs and per-node MIPs on other MEGs for which it is a transit node.

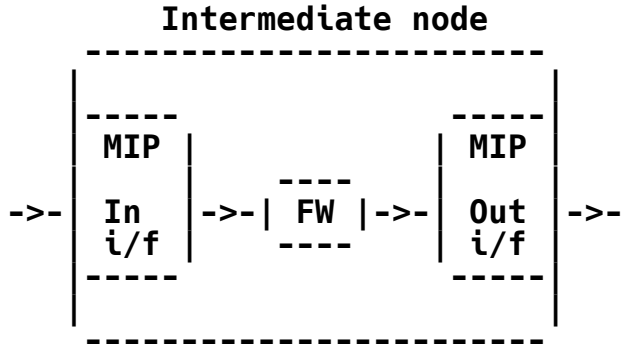


Figure 4: Example of Per-Interface MIPs

Figure 4 describes an example of two per-interface MIPs at an intermediate node of a point-to-point MEG.

Using per-interface MIPs allows the network operator to determine that the location of a failure or performance degradation is within a node or on a link between two adjacent nodes.

When sending an OAM packet to a MIP, the source MEP should set the TTL field to indicate the number of hops necessary to reach the node where the MIP resides.

The source MEP should also include target MIP information in the OAM packets sent to a MIP to allow proper identification of the MIP within the node. The MEG the OAM packet belongs to is associated with the MPLS label, whether the label is used to infer the MEG or the content of the OAM packet is an implementation choice. In the latter case, the MPLS label is checked to be the expected one.

The use of TTL expiry to deliver OAM packets to a specific MIP is not a fully reliable delivery mechanism because the TTL distance of a MIP from a MEP can change. Any MPLS-TP node silently discards any OAM packet that is received with an expired TTL and that is not addressed to any of its MIPs or MEPs. An MPLS-TP node that does not support OAM is also expected to silently discard any received OAM packet.

Packets directed to a MIP may not necessarily carry specific MIP identification information beyond that of TTL distance. In this case, a MIP would promiscuously respond to all MEP queries on its MEG. This capability could be used for discovery functions (e.g., route tracing as defined in Section 6.4) or when it is desirable to leave to the originating MEP the job of correlating TTL and MIP identifiers and noting changes or irregularities (via comparison with information previously extracted from the network).

MIPs are associated to the MEG they belong to, and their identity is unique within the MEG. However, their identity is not necessarily unique to the MEG, e.g., all nodal MIPs in a node can have a common identity.

A node hosting a MEP can also support per-interface Up MEPs and per-interface MIPs on either side of the forwarding engine.

Once a MEG is configured, the operator can enable/disable the MIPs on the nodes within the MEG. All the intermediate nodes and possibly the end nodes host MIP(s). Local policy allows them to be enabled per function and per MEG. The local policy is controlled by the management system, which may delegate it to the control plane. A disabled MIP silently discards any received OAM packets.

3.5. Server MEPs

A server MEP is a MEP of a MEG that is either:

- defined in a layer network that is "below", which is to say encapsulates and transports the MPLS-TP layer network being referenced; or
- defined in a sub-layer of the MPLS-TP layer network that is "below", which is to say encapsulates and transports the sub-layer being referenced.

A server MEP can coincide with a MIP or a MEP in the client (MPLS-TP) (sub-)layer network.

A server MEP also provides server-layer OAM indications to the client/server adaptation function between the client (MPLS-TP) (sub-)layer network and the server (sub-)layer network. The adaptation function maintains state on the mapping of MPLS-TP transport paths that are set up over that server (sub-)layer's transport path.

For example, a server MEP can be:

- a non-MPLS MEP at a termination point of a physical link (e.g., 802.3, an SDH Virtual Circuit, or OTN Optical Data Unit (ODU)), for the MPLS-TP Section layer network, defined in Section 4.1;
- an MPLS-TP Section MEP for MPLS-TP LSPs, defined in Section 4.2;
- an MPLS-TP LSP MEP for MPLS-TP PWs, defined in Section 4.3;

- an MPLS-TP SPME MEP used for LSP path segment monitoring, as defined in Section 4.4, for MPLS-TP LSPs or higher-level SPMEs providing LSP path segment monitoring; or

- an MPLS-TP SPME MEP used for PW path segment monitoring, as defined in Section 4.5, for MPLS-TP PWs or higher-level SPMEs providing PW path segment monitoring.

The server MEP can run appropriate OAM functions for fault detection within the server (sub-)layer network and provides a fault indication to its client MPLS-TP layer network via the client/server adaptation function. When the server layer is not MPLS-TP, server MEP OAM functions are simply assumed to exist but are outside the scope of this document.

3.6. Configuration Considerations

When a control plane is not present, the management plane configures these functional components. Otherwise, they can be configured by either the management plane or the control plane.

Local policy allows disabling the usage of any available "out-of-band" return path, as defined in [8], irrespective of what is requested by the node originating the OAM packet.

SPMEs are usually instantiated when the transport path is created by either the management plane or the control plane (if present). Sometimes an SPME can be instantiated after the transport path is initially created.

3.7. P2MP Considerations

All the traffic sent over a P2MP transport path, including OAM packets generated by a MEP, is sent (multicast) from the root to all the leaves. As a consequence:

- To send an OAM packet to all leaves, the source MEP can send a single OAM packet that will be delivered by the forwarding plane to all the leaves and processed by all the leaves. Hence, a single OAM packet can simultaneously instrument all the MEs in a P2MP MEG.

- To send an OAM packet to a single leaf, the source MEP sends a single OAM packet that will be delivered by the forwarding plane to all the leaves but contains sufficient information to identify a target leaf, and therefore is processed only by the target leaf and can be silently discarded by the other leaves.

- To send an OAM packet to a single MIP, the source MEP sends a single OAM packet with the TTL field indicating the number of hops necessary to reach the node where the MIP resides. This packet will be delivered by the forwarding plane to all intermediate nodes at the same TTL distance of the target MIP and to any leaf that is located at a shorter distance. The OAM packet must contain sufficient information to identify the target MIP and therefore is processed only by the target MIP and can be silently discarded by the others.

- In order to send an OAM packet to M leaves (i.e., a subset of all the leaves), the source MEP sends M different OAM packets targeted to each individual leaf in the group of M leaves. Aggregating or subsetting mechanisms are outside the scope of this document.

A bud node with a Down MEP or a per-node MEP will both terminate and relay OAM packets. Similar to how fault coverage is maximized by the explicit utilization of Up MEPs, the same is true for MEPs on a bud node.

P2MP paths are unidirectional; therefore, any return path to an originating MEP for on-demand transactions will be out-of-band. A mechanism to target "on-demand" transactions to a single MEP or MIP is required as it relieves the originating MEP of an arbitrarily large processing load and of the requirement to filter and discard undesired responses. This is because normally TTL exhaustion will address all MIPs at a given distance from the source, and failure to exhaust TTL will address all MEPs.

3.8. Further Considerations of Enhanced Segment Monitoring

Segment monitoring, like any in-service monitoring, in a transport network should meet the following network objectives:

1. The monitoring and maintenance of existing transport paths has to be conducted in service without traffic disruption.

2. Segment monitoring must not modify the forwarding of the segment portion of the transport path.

SPMEs defined in Section 3.2 meet the above two objectives, when they are pre-configured or pre-instantiated as exemplified in Section 3.6. However, sometimes pre-design and pre-configuration of all the considered patterns of SPME are not preferable in real operation due to the burden of design works, a number of header consumptions, bandwidth consumption, and so on.

When SPMEs are configured or instantiated after the transport path has been created, network objective (1) can be met: application and removal of SPME to a faultless monitored transport entity can be performed in such a way as not to introduce any loss of traffic, e.g., by using a non-disruptive "make before break" technique.

However, network objective (2) cannot be met due to new assignment of MPLS labels. As a consequence, generally speaking, the results of SPME monitoring are not necessarily correlated with the behavior of traffic in the monitored entity when it does not use SPME. For example, application of SPME to a problematic/faulty monitoring entity might "fix" the problem encountered by the latter -- for as long as SPME is applied. And vice versa, application of SPME to a faultless monitored entity may result in making it faulty -- again, as long as SPME is applied.

Support for a more sophisticated segment-monitoring mechanism (temporal and hitless segment monitoring) to efficiently meet the two network objectives may be necessary.

One possible option to instantiate non-intrusive segment monitoring without the use of SPMEs would require the MIPs selected as monitoring end points to implement enhanced functionality and state for the monitored transport path.

For example, the MIPs need to be configured with the TTL distance to the peer or with the address of the peer, when out-of-band return paths are used.

A further issue that would need to be considered is events that result in changing the TTL distance to the peer monitoring entity, such as protection events that may temporarily invalidate OAM information gleaned from the use of this technique.

Further considerations on this technique are outside the scope of this document.

4. Reference Model

The reference model for the MPLS-TP OAM framework builds upon the concept of a MEG, and its associated MEPs and MIPs, to support the functional requirements specified in RFC 5860 [11].

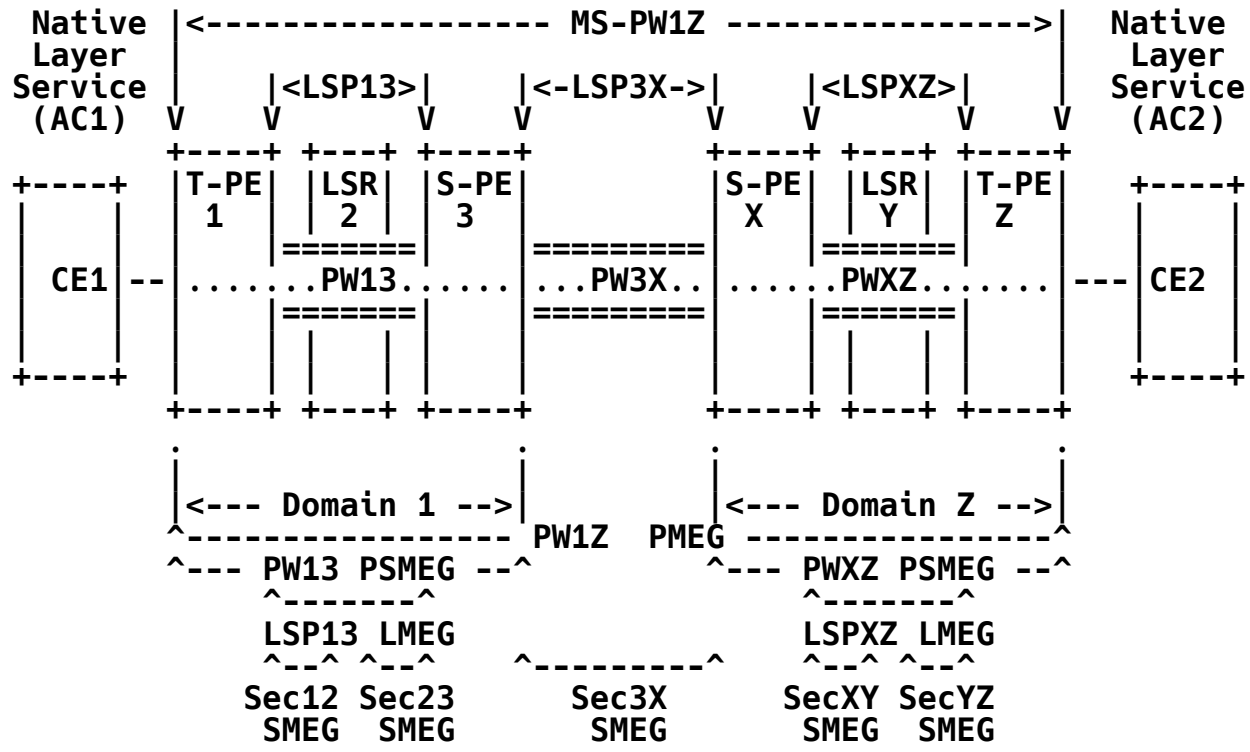
The following MPLS-TP MEGs are specified in this document:

- A Section Maintenance Entity Group (SMEG), allowing monitoring and management of MPLS-TP Sections (between MPLS LSRs).

- An LSP Maintenance Entity Group (LMEG), allowing monitoring and management of an end-to-end LSP (between LERs).
- A PW Maintenance Entity Group (PMEG), allowing monitoring and management of an end-to-end Single-Segment Pseudowire (SS-PW) or MS-PW (between T-PEs).
- An LSP SPME ME Group (LSMEG), allowing monitoring and management of an SPME (between a given pair of LERs and/or LSRs along an LSP).
- A PW SPME ME Group (PSMEG), allowing monitoring and management of an SPME (between a given pair of T-PEs and/or S-PEs along an (MS-)PW).

The MEGs specified in this MPLS-TP OAM framework are compliant with the architecture framework for MPLS-TP [8] that includes both MS-PWs [4] and LSPs [1].

Hierarchical LSPs are also supported in the form of SPMEs. In this case, each LSP in the hierarchy is a different sub-layer network that can be monitored, independently from higher- and lower-level LSPs in the hierarchy, on an end-to-end basis (from LER to LER) by an SPME. It is possible to monitor a portion of a hierarchical LSP by instantiating a hierarchical SPME between any LERs/LSRs along the hierarchical LSP.



^---^ ME
^ MEP
==== LSP
.... PW

T-PE 1: Terminating Provider Edge 1
LSR 2: Label Switching Router 2
S-PE 3: Switching Provider Edge 3
S-PE X: Switching Provider Edge X
LSR Y: Label Switching Router Y
T-PE Z: Terminating Provider Edge Z

Figure 5: Reference Model for the MPLS-TP OAM Framework

Figure 5 depicts a high-level reference model for the MPLS-TP OAM framework. The figure depicts portions of two MPLS-TP-enabled network domains, Domain 1 and Domain Z. In Domain 1, T-PE 1 is adjacent to LSR 2 via the MPLS-TP Section Sec12, and LSR 2 is adjacent to S-PE 3 via the MPLS-TP Section Sec23. Similarly, in Domain Z, S-PE X is adjacent to LSR Y via the MPLS-TP Section SecXY, and LSR Y is adjacent to T-PE Z via the MPLS-TP Section SecYZ. In addition, S-PE 3 is adjacent to S-PE X via the MPLS-TP Section Sec3X.

Figure 5 also shows a bidirectional MS-PW (MS-PW1Z) between AC1 on T-PE1 and AC2 on T-PE Z. The MS-PW consists of three bidirectional PW path segments: 1) PW13 path segment between T-PE 1 and S-PE 3 via the bidirectional LSP13 LSP, 2) PW3X path segment between S-PE 3 and S-PE X via the bidirectional LSP3X LSP, and 3) PWXZ path segment between S-PE X and T-PE Z via the bidirectional LSPXZ LSP.

The MPLS-TP OAM procedures that apply to a MEG are expected to operate independently from procedures on other MEGs. Yet, this does not preclude that multiple MEGs may be affected simultaneously by the same network condition -- for example, a fiber cut event.

Note that there are no constraints imposed by this OAM framework on the number or type (P2P, P2MP, LSP, or PW), of MEGs that may be instantiated on a particular node. In particular, when looking at Figure 5, it should be possible to configure one or more MEPs on the same node if that node is the end point of one or more MEGs.

Figure 5 does not describe a PW3X PSMEG because typically SPMEs are used to monitor an OAM domain (like PW13 and PWXZ PSMEGs) rather than the segment between two OAM domains. However, the OAM framework does not pose any constraints on the way SPMEs are instantiated as long as they are not overlapping.

The subsections below define the MEGs specified in this MPLS-TP OAM architecture framework document. Unless otherwise stated, all references to domains, LSRs, MPLS-TP Sections, LSPs, pseudowires, and MEGs in this section are made in relation to those shown in Figure 5.

4.1. MPLS-TP Section Monitoring (SMEG)

An MPLS-TP Section MEG (SMEG) is an MPLS-TP maintenance entity intended to monitor an MPLS-TP Section. An SMEG may be configured on any MPLS-TP section. SMEG OAM packets must fate-share with the user data packets sent over the monitored MPLS-TP Section.

An SMEG is intended to be deployed for applications where it is preferable to monitor the link between topologically adjacent (next hop in this layer network) MPLS-TP LSRs rather than monitoring the individual LSP or PW path segments traversing the MPLS-TP Section and where the server-layer technology does not provide adequate OAM capabilities.

Figure 5 shows five Section MEGs configured in the network between AC1 and AC2:

1. Sec12 MEG associated with the MPLS-TP Section between T-PE 1 and LSR 2,

2. Sec23 MEG associated with the MPLS-TP Section between LSR 2 and S-PE 3,

3. Sec3X MEG associated with the MPLS-TP Section between S-PE 3 and S-PE X,

4. SecXY MEG associated with the MPLS-TP Section between S-PE X and LSR Y, and

5. SecYZ MEG associated with the MPLS-TP Section between LSR Y and T-PE Z

4.2. MPLS-TP LSP End-to-End Monitoring Group (LMEG)

An MPLS-TP LSP MEG (LMEG) is an MPLS-TP maintenance entity group intended to monitor an end-to-end LSP between its LERs. An LMEG may be configured on any MPLS LSP. LMEG OAM packets must fate-share with user data packets sent over the monitored MPLS-TP LSP.

An LMEG is intended to be deployed in scenarios where it is desirable to monitor an entire LSP between its LERs, rather than, say, monitoring individual PWs.

Figure 5 depicts two LMEGs configured in the network between AC1 and AC2: 1) the LSP13 LMEG between T-PE 1 and S-PE 3, and 2) the LSPXZ LMEG between S-PE X and T-PE Z. Note that the presence of a LSP3X LMEG in such a configuration is optional, and hence, not precluded by this framework. For instance, the network operator may prefer to monitor the MPLS-TP Section between the two LSRs rather than the individual LSPs.

4.3. MPLS-TP PW Monitoring (PMEG)

An MPLS-TP PW MEG (PMEG) is an MPLS-TP maintenance entity intended to monitor a SS-PW or MS-PW between its T-PEs. A PMEG can be configured on any SS-PW or MS-PW. PMEG OAM packets must fate-share with the user data packets sent over the monitored PW.

A PMEG is intended to be deployed in scenarios where it is desirable to monitor an entire PW between a pair of MPLS-TP-enabled T-PEs rather than monitoring the LSP that aggregates multiple PWs between PEs.

Figure 5 depicts an MS-PW (MS-PW1Z) consisting of three path segments (PW13, PW3X, and PWXZ) and its associated end-to-end PMEG (PW1Z PMEG).

4.4. MPLS-TP LSP SPME Monitoring (LSMEG)

An MPLS-TP LSP SPME MEG (LSMEG) is an MPLS-TP SPME with an associated maintenance entity group intended to monitor an arbitrary part of an LSP between the MEPs instantiated for the SPME, independent from the end-to-end monitoring (LMEG). An LSMEG can monitor an LSP path segment, and it may also include the forwarding engine(s) of the node(s) at the edge(s) of the path segment.

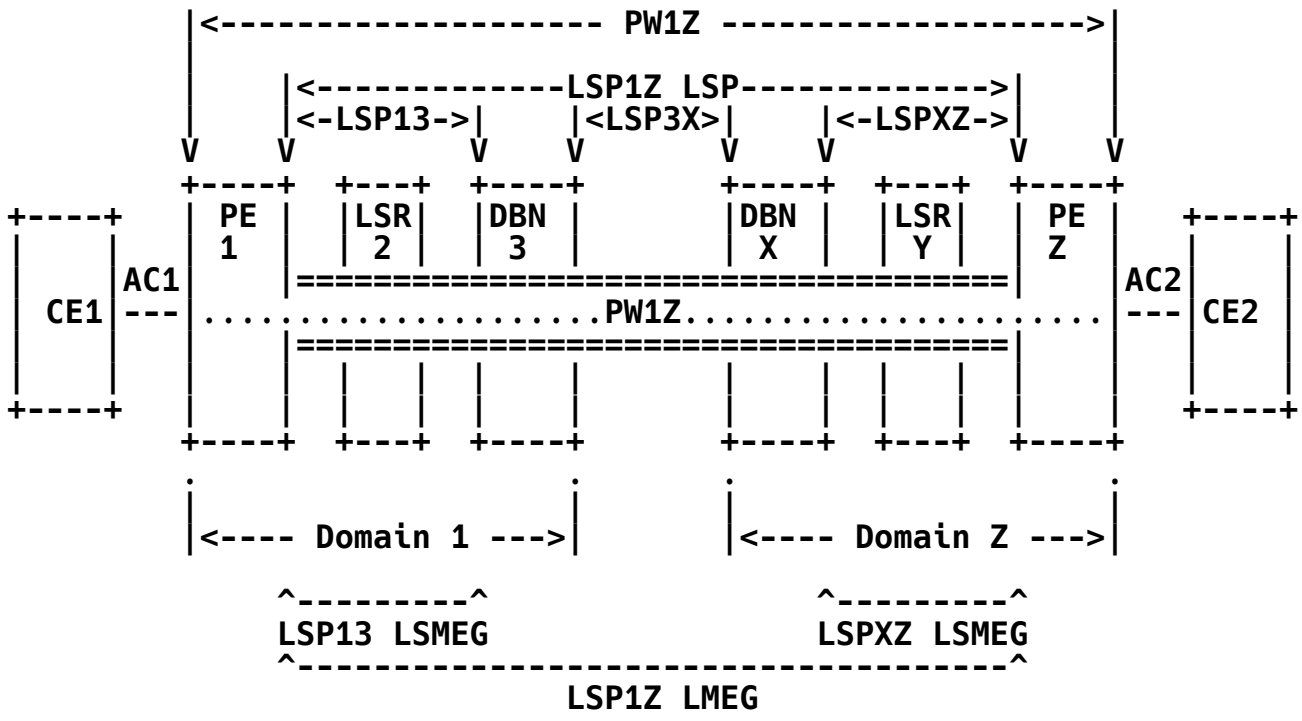
When an SPME is established between non-adjacent LSRs, the edges of the SPME become adjacent at the LSP sub-layer network and any LSR that was previously in between becomes an LSR for the SPME.

Multiple hierarchical LSMEGs can be configured on any LSP. LSMEG OAM packets must fate-share with the user data packets sent over the monitored LSP path segment.

A LSME can be defined between the following entities:

- The LER and LSR of a given LSP.
- Any two LSRs of a given LSP.

An LSMEG is intended to be deployed in scenarios where it is preferable to monitor the behavior of a part of an LSP or set of LSPs rather than the entire LSP itself, for example, when there is a need to monitor a part of an LSP that extends beyond the administrative boundaries of an MPLS-TP-enabled administrative domain.



DBN: Domain Border Node

PE 1: Provider Edge 1
LSR 2: Label Switching Router 2
DBN 3: Domain Border Node 3
DBN X: Domain Border Node X
LSR Y: Label Switching Router Y
PE Z: Provider Edge Z

Figure 6: MPLS-TP LSP SPME MEG (LSMEG)

Figure 6 depicts a variation of the reference model in Figure 5 where there is an end-to-end LSP (LSP1Z) between PE 1 and PE Z. LSP1Z consists of, at least, three LSP Concatenated Segments: LSP13, LSP3X, and LSPXZ. In this scenario, there are two separate LSMEGs configured to monitor the LSP1Z: 1) a LSMEG monitoring the LSP13 Concatenated Segment on Domain 1 (LSP13 LSMEG), and 2) a LSMEG monitoring the LSPXZ Concatenated Segment on Domain Z (LSPXZ LSMEG).

It is worth noticing that LSMEGs can coexist with the LMEG monitoring the end-to-end LSP and that LSMEG MEPs and LMEG MEPs can be coincident in the same node (e.g., PE 1 node supports both the LSP1Z LMEG MEP and the LSP13 LSMEG MEP).

4.5. MPLS-TP MS-PW SPME Monitoring (PSMEG)

An MPLS-TP MS-PW SPME Monitoring MEG (PSMEG) is an MPLS-TP SPME with an associated maintenance entity group intended to monitor an arbitrary part of an MS-PW between the MEPs instantiated for the SPME, independently of the end-to-end monitoring (PMEG). A PSMEG can monitor a PW path segment, and it may also include the forwarding engine(s) of the node(s) at the edge(s) of the path segment. A PSMEG is no different than an SPME; it is simply named as such to discuss SPMEs specifically in a PW context.

When SPME is established between non-adjacent S-PEs, the edges of the SPME become adjacent at the MS-PW sub-layer network, and any S-PE that was previously in between becomes an LSR for the SPME.

S-PE placement is typically dictated by considerations other than OAM. S-PEs will frequently reside at operational boundaries such as the transition from distributed control plane (CP) to centralized Network Management System (NMS) control or at a routing area boundary. As such, the architecture would appear not to have the flexibility that arbitrary placement of SPME segments would imply. Support for an arbitrary placement of PSMEG would require the definition of additional PW sub-layering. Multiple hierarchical PSMEGs can be configured on any MS-PW. PSMEG OAM packets fate-share with the user data packets sent over the monitored PW path Segment.

A PSMEG does not add hierarchical components to the MPLS architecture; it defines the role of existing components for the purposes of discussing OAM functionality.

A PSME can be defined between the following entities:

- The T-PE and any S-PE of a given MS-PW.
- Any two S-PEs of a given MS-PW.

Note that, in line with the SPME description in Section 3.2, when a PW SPME is instantiated after the MS-PW has been instantiated, the TTL distance of the MIPs may change and MIPs in the PW SPME are no longer part of the encompassing MEG. This means that the S-PE nodes hosting these MIPs are no longer S-PEs but P nodes at the SPME LSP level. The consequences are that the S-PEs hosting the PSMEG MEPs become adjacent S-PEs. This is no different than the operation of SPMEs in general.

A PSMEG is intended to be deployed in scenarios where it is preferable to monitor the behavior of a part of an MS-PW rather than the entire end-to-end PW itself, for example, when monitoring an MS-

PW path segment within a given network domain of an inter-domain MS-PW.

Figure 5 depicts an MS-PW (MS-PW1Z) consisting of three path segments: PW13, PW3X, and PWXZ with two separate PSMEGs: 1) a PSMEG monitoring the PW13 MS-PW path segment on Domain 1 (PW13 PSMEG) and 2) a PSMEG monitoring the PWXZ MS-PW path segment on Domain Z with (PWXZ PSMEG).

It is worth noticing that PSMEGs can coexist with the PMEG monitoring the end-to-end MS-PW and that PSMEG MEPs and PMEG MEPs can be coincident in the same node (e.g., T-PE 1 node supports both the PW1Z PMEG MEP and the PW13 PSMEG MEP).

4.6. Fate-Sharing Considerations for Multilink

Multilink techniques are in use today and are expected to continue to be used in future deployments. These techniques include Ethernet link aggregation [22] and the use of link bundling for MPLS [18] where the option to spread traffic over component links is supported and enabled. While the use of link bundling can be controlled at the MPLS-TP layer, use of link aggregation (or any server-layer-specific multilink) is not necessarily under the control of the MPLS-TP layer. Other techniques may emerge in the future. These techniques frequently share the characteristic that an LSP may be spread over a set of component links and therefore be reordered, but no flow within the LSP is reordered (except when very infrequent and minimally disruptive load rebalancing occurs).

The use of multilink techniques may be prohibited or permitted in any particular deployment. If multilink techniques are used, the deployment can be considered to be only partially MPLS-TP compliant; however, this is unlikely to prevent their use.

The implications for OAM are that not all components of a multilink will be exercised, independent server-layer OAM being required to exercise the aggregated link components. This has further implications for MIP and MEP placement, as per-interface MIPs or Down MEPs on a multilink interface are akin to a layer violation, as they instrument at the granularity of the server layer. The implications for reduced OAM loss measurement functionality are documented in Sections 5.5.3 and 6.2.3.

5. OAM Functions for Proactive Monitoring

In this document, proactive monitoring refers to OAM operations that are either configured to be carried out periodically and continuously or preconfigured to act on certain events such as alarm signals.

Proactive monitoring is usually performed "in-service". Such transactions are universally MEP to MEP in operation, while notifications can be node to node (e.g., some MS-PW transactions) or node to MEPs (e.g., AIS). The control and measurement considerations are:

1. Proactive monitoring for a MEG is typically configured at the creation time of the transport path.

2. The operational characteristics of in-band measurement transactions (e.g., CV, Loss Measurement (LM), etc.) are configured at the MEPs.

3. Server-layer events are reported by OAM packets originating at intermediate nodes.

4. The measurements resulting from proactive monitoring are typically reported outside of the MEG (e.g., to a management system) as notification events such as faults or indications of performance degradations (such as signal degrade conditions).

5. The measurements resulting from proactive monitoring may be periodically harvested by an NMS.

Proactive fault reporting is assumed to be subject to unreliable delivery and soft-state, and it needs to operate in cases where a return path is not available or faulty. Therefore, periodic repetition is assumed to be used for reliability, instead of handshaking.

Delay measurement also requires periodic repetition to allow estimation of the packet delay variation for the MEG.

For statically provisioned transport paths, the above information is statically configured; for dynamically established transport paths, the configuration information is signaled via the control plane or configured via the management plane.

The operator may enable/disable some of the consequent actions defined in Section 5.1.2.

5.1. Continuity Check and Connectivity Verification

Proactive Continuity Check functions, as required in Section 2.2.2 of RFC 5860 [11], are used to detect a loss of continuity (LOC) defect between two MEPs in a MEG.

Proactive Connectivity Verification functions, as required in Section 2.2.3 of RFC 5860 [11], are used to detect an unexpected connectivity defect between two MEGs (e.g., mismerging or misconnection), as well as unexpected connectivity within the MEG with an unexpected MEP.

Both functions are based on the (proactive) generation, at the same rate, of OAM packets by the source MEP that are processed by the peer sink MEP(s). As a consequence, in order to save OAM bandwidth consumption, CV, when used, is linked with CC into Continuity Check and Connectivity Verification (CC-V) OAM packets.

In order to perform proactive Connectivity Verification, each CC-V OAM packet also includes a globally unique Source MEP identifier, whose value needs to be configured on the source MEP and on the peer sink MEP(s). In some cases, to avoid the need to configure the globally unique Source MEP identifier, it is preferable to perform only proactive Continuity Check. In this case, the CC-V OAM packet does not need to include any globally unique Source MEP identifier. Therefore, a MEG can be monitored only for CC or for both CC and CV. CC-V OAM packets used for CC-only monitoring are called CC OAM packets, while CC-V OAM packets used for both CC and CV are called CV OAM packets.

As a consequence, it is not possible to detect misconnections between two MEGs monitored only for continuity as neither the OAM packet type nor the OAM packet content provides sufficient information to disambiguate an invalid source. To expand:

- For a CC OAM packet leaking into a CC monitored MEG - undetectable.
- For a CV OAM packet leaking into a CC monitored MEG - reception of CV OAM packets instead of a CC OAM packets (e.g., with the additional Source MEP identifier) allows detecting the fault.
- For a CC OAM packet leaking into a CV monitored MEG - reception of CC OAM packets instead of CV OAM packets (e.g., lack of additional Source MEP identifier) allows detecting the fault.
- For a CV OAM packet leaking into a CV monitored MEG - reception of CV OAM packets with different Source MEP identifier permits fault to be identified.

Having a common packet format for CC-V OAM packets would simplify parsing in a sink MEP to properly detect all the misconfiguration cases described above.

MPLS-TP OAM supports different formats of MEP identifiers to address different environments. When an alternative to IP addressing is desired (e.g., MPLS-TP is deployed in transport network environments where consistent operations with other transport technologies defined by the ITU-T are required), the ITU Carrier Code (ICC)-based format for MEP identification is used: this format is under definition in [25]. When MPLS-TP is deployed in an environment where IP capabilities are available and desired for OAM, the IP-based MEP identification is used: this format is described in [24].

CC-V OAM packets are transmitted at a regular, operator-configurable rate. The default CC-V transmission periods are application dependent (see Section 5.1.3).

Proactive CC-V OAM packets are transmitted with the "minimum loss probability PHB" within the transport path (LSP, PW) they are monitoring. For E-LSPs, this PHB is configurable on the network operator's basis, while for L-LSPs this is determined as per RFC 3270 [23]. PHBs can be translated at the network borders by the same function that translates them for user data traffic. The implication is that CC-V fate-shares with much of the forwarding implementation, but not all aspects of PHB processing are exercised. Either on-demand tools are used for finer-grained fault finding or an implementation may utilize a CC-V flow per PHB to ensure a CC-V flow fate-shares with each individual PHB.

In a co-routed or associated, bidirectional point-to-point transport path, when a MEP is enabled to generate proactive CC-V OAM packets with a configured transmission rate, it also expects to receive proactive CC-V OAM packets from its peer MEP at the same transmission rate. This is because a common SLA applies to all components of the transport path. In a unidirectional transport path (either point-to-point or point-to-multipoint), the source MEP is enabled only to generate CC-V OAM packets, while each sink MEP is configured to expect these packets at the configured rate.

MIPs, as well as intermediate nodes not supporting MPLS-TP OAM, are transparent to the proactive CC-V information and forward these proactive CC-V OAM packets as regular data packets.

During path setup and tear down, situations arise where CC-V checks would give rise to alarms, as the path is not fully instantiated. In order to avoid these spurious alarms, the following procedures are recommended. At initialization, the source MEP function (generating

proactive CC-V packets) should be enabled prior to the corresponding sink MEP function (detecting continuity and connectivity defects). When disabling the CC-V proactive functionality, the sink MEP function should be disabled prior to the corresponding source MEP function.

It should be noted that different encapsulations are possible for CC-V packets, and therefore it is possible that in case of misconfigurations or mis-connectivity, CC-V packets are received with an unexpected encapsulation.

There are practical limitations to detecting unexpected encapsulation. It is possible that there are misconfiguration or mis-connectivity scenarios where OAM packets can alias as payload, e.g., when a transport path can carry an arbitrary payload without a pseudowire.

When CC-V packets are received with an unexpected encapsulation that can be parsed by a sink MEP, the CC-V packet is processed as if it were received with the correct encapsulation. If it is not a manifestation of a mis-connectivity defect, a warning is raised (see Section 5.1.1.4). Otherwise, the CC-V packet may be silently discarded as unrecognized and a LOC defect may be detected (see Section 5.1.1.1).

The defect conditions are described in no specific order.

5.1.1. Defects Identified by CC-V

Proactive CC-V functions allow a sink MEP to detect the defect conditions described in the following subsections. For all of the described defect cases, a sink MEP should notify the equipment fault management process of the detected defect.

Sequential consecutive loss of CC-V packets is considered indicative of an actual break and not of congestive loss or physical-layer degradation. The loss of 3 packets in a row (implying a detection interval that is 3.5 times the insertion time) is interpreted as a true break and a condition that will not clear by itself.

A CC-V OAM packet is considered to carry an unexpected globally unique Source MEP identifier if it is a CC OAM packet received by a sink MEP monitoring the MEG for CV; it is a CV OAM packet received by a sink MEP monitoring the MEG for CC, or it is a CV OAM packet received by a sink MEP monitoring the MEG for CV but carrying a unique Source MEP identifier that is different that the expected one. Conversely, the CC-V packet is considered to have an expected globally unique Source MEP identifier; it is a CC OAM packet received

by a sink MEP monitoring the MEG for CC, or it is a CV OAM packet received by a sink MEP monitoring the MEG for CV and carrying a unique Source MEP identifier that is equal to the expected one.

5.1.1.1. Loss of Continuity Defect

When proactive CC-V is enabled, a sink MEP detects a loss of continuity (LOC) defect when it fails to receive proactive CC-V OAM packets from the source MEP.

- Entry criteria: If no proactive CC-V OAM packets from the source MEP (and in the case of CV, this includes the requirement to have the expected globally unique Source MEP identifier) are received within the interval equal to 3.5 times the receiving MEP's configured CC-V reception period.

- Exit criteria: A proactive CC-V OAM packet from the source MEP (and again in the case of CV, with the expected globally unique Source MEP identifier) is received.

5.1.1.2. Mis-Connectivity Defect

When a proactive CC-V OAM packet is received, a sink MEP identifies a mis-connectivity defect (e.g., mismerge, misconnection, or unintended looping) when the received packet carries an unexpected globally unique Source MEP identifier.

- Entry criteria: The sink MEP receives a proactive CC-V OAM packet with an unexpected globally unique Source MEP identifier or with an unexpected encapsulation.

- Exit criteria: The sink MEP does not receive any proactive CC-V OAM packet with an unexpected globally unique Source MEP identifier for an interval equal at least to 3.5 times the longest transmission period of the proactive CC-V OAM packets received with an unexpected globally unique Source MEP identifier since this defect has been raised. This requires the OAM packet to self-identify the CC-V periodicity, as not all MEPs can be expected to have knowledge of all MEGs.

5.1.1.3. Period Misconfiguration Defect

If proactive CC-V OAM packets are received with the expected globally unique Source MEP identifier but with a transmission period different than the locally configured reception period, then a CC-V period misconfiguration defect is detected.

- Entry criteria: A MEP receives a CC-V proactive packet with the expected globally unique Source MEP identifier but with a transmission period different than its own CC-V-configured transmission period.

- Exit criteria: The sink MEP does not receive any proactive CC-V OAM packet with the expected globally unique Source MEP identifier and an incorrect transmission period for an interval equal at least to 3.5 times the longest transmission period of the proactive CC-V OAM packets received with the expected globally unique Source MEP identifier and an incorrect transmission period since this defect has been raised.

5.1.1.4. Unexpected Encapsulation Defect

If proactive CC-V OAM packets are received with the expected globally unique Source MEP identifier but with an unexpected encapsulation, then a CC-V unexpected encapsulation defect is detected.

It should be noted that there are practical limitations to detecting unexpected encapsulation (see Section 5.1.1).

- Entry criteria: A MEP receives a CC-V proactive packet with the expected globally unique Source MEP identifier but with an unexpected encapsulation.

- Exit criteria: The sink MEP does not receive any proactive CC-V OAM packet with the expected globally unique Source MEP identifier and an unexpected encapsulation for an interval equal at least to 3.5 times the longest transmission period of the proactive CC-V OAM packets received with the expected globally unique Source MEP identifier and an unexpected encapsulation since this defect has been raised.

5.1.2. Consequent Action

A sink MEP that detects any of the defect conditions defined in Section 5.1.1 declares a defect condition and performs the following consequent actions.

If a MEP detects a mis-connectivity defect, it blocks all the traffic (including also the user data packets) that it receives from the misconnected transport path.

If a MEP detects a LOC defect that is not caused by a period misconfiguration, it should block all the traffic (including also the user data packets) that it receives from the transport path, if this consequent action has been enabled by the operator.

It is worth noticing that the OAM requirements document [11] recommends that CC-V proactive monitoring be enabled on every MEG in order to reliably detect connectivity defects. However, CC-V proactive monitoring can be disabled by an operator for a MEG. In the event of a misconnection between a transport path that is proactively monitored for CC-V and a transport path that is not, the MEP of the former transport path will detect a LOC defect representing a connectivity problem (e.g., a misconnection with a transport path where CC-V proactive monitoring is not enabled) instead of a continuity problem, with a consequence of delivery of traffic to an incorrect destination. For these reasons, the traffic block consequent action is applied even when a LOC condition occurs. This block consequent action can be disabled through configuration. This deactivation of the block action may be used for activating or deactivating the monitoring when it is not possible to synchronize the function activation of the two peer MEPs.

If a MEP detects a LOC defect (Section 5.1.1.1) or a mis-connectivity defect (Section 5.1.1.2), it declares a signal fail condition of the ME.

It is a matter of local policy whether or not a MEP that detects a period misconfiguration defect (Section 5.1.1.3) declares a signal fail condition of the ME.

The detection of an unexpected encapsulation defect does not have any consequent action: it is just a warning for the network operator. An implementation able to detect an unexpected encapsulation but not able to verify the source MEP ID may choose to declare a mis-connectivity defect.

5.1.3. Configuration Considerations

At all MEPs inside a MEG, the following configuration information needs to be configured when a proactive CC-V function is enabled:

- MEG-ID: the MEG identifier to which the MEP belongs.
- MEP-ID: the MEP's own identity inside the MEG.
- list of the other MEPs in the MEG. For a point-to-point MEG, the list would consist of the single MEP ID from which the OAM packets are expected. In case of the root MEP of a P2MP MEG, the list is composed of all the leaf MEP IDs inside the MEG. In case of the leaf MEP of a P2MP MEG, the list is composed of the root MEP ID (i.e., each leaf needs to know the root MEP ID from which it expects to receive the CC-V OAM packets).

- PHB for E-LSPs. It identifies the per-hop behavior of a CC-V packet. Proactive CC-V packets are transmitted with the "minimum loss probability PHB" previously configured within a single network operator. This PHB is configurable on network operator's basis. PHBs can be translated at the network borders.

- transmission rate. The default CC-V transmission periods are application dependent (depending on whether they are used to support fault management, performance monitoring, or protection-switching applications):

  * Fault Management: default transmission period is 1 s (i.e., transmission rate of 1 packet/second).

  * Performance Management: default transmission period is 100 ms (i.e., transmission rate of 10 packets/second). CC-V contributes to the accuracy of performance monitoring statistics by permitting the defect-free periods to be properly distinguished as described in Sections 5.5.1 and 5.6.1.

  * Protection Switching: If protection switching with CC-V, defect entry criteria of 12 ms is required (for example, in conjunction with the requirement to support 50 ms recovery time as indicated in RFC 5654 [5]), then an implementation should use a default transmission period of 3.33 ms (i.e., transmission rate of 300 packets/second). Sometimes, the requirement of 50 ms recovery time is associated with the requirement for a CC-V defect entry criteria period of 35 ms; in these cases a transmission period of 10 ms (i.e., transmission rate of 100 packets/second) can be used. Furthermore, when there is no need for so small CC-V defect entry criteria periods, a larger transmission period can be used.

  It should be possible for the operator to configure these transmission rates for all applications, to satisfy specific network requirements.

  Note that the reception period is the same as the configured transmission rate.

  For management-provisioned transport paths, the above parameters are statically configured; for dynamically signaled transport paths, the configuration information is distributed via the control plane.

  The operator should be able to enable/disable some of the consequent actions. Which consequent actions can be enabled/disabled is described in Section 5.1.2.

5.2. Remote Defect Indication

The Remote Defect Indication (RDI) function, as required in Section 2.2.9 of RFC 5860 [11], is an indicator that is transmitted by a sink MEP to communicate to its source MEP that a signal fail condition exists. In case of co-routed and associated bidirectional transport paths, RDI is associated with proactive CC-V, and the RDI indicator can be piggy-backed onto the CC-V packet. In case of unidirectional transport paths, the RDI indicator can be sent only using an out-of-band return path if it exists and its usage is enabled by policy actions.

When a MEP detects a signal fail condition (e.g., in case of a continuity or connectivity defect), it should begin transmitting an RDI indicator to its peer MEP. When incorporated into CC-V, the RDI information will be included in all proactive CC-V packets that it generates for the duration of the signal fail condition's existence.

A MEP that receives packets from a peer MEP with the RDI information should determine that its peer MEP has encountered a defect condition associated with a signal fail condition.

MIPs as well as intermediate nodes not supporting MPLS-TP OAM are transparent to the RDI indicator and forward OAM packets that include the RDI indicator as regular data packets, i.e., the MIP should not perform any actions nor examine the indicator.

When the signal fail condition clears, the MEP should stop transmitting the RDI indicator to its peer MEP. When incorporated into CC-V, the RDI indicator will not be set for subsequent transmission of proactive CC-V packets. A MEP should clear the RDI defect upon reception of an RDI indicator cleared.

5.2.1. Configuration Considerations

In order to support RDI, the indication may be carried in a unique OAM packet or may be embedded in a CC-V packet. The in-band RDI transmission rate and PHB of the OAM packets carrying RDIs should be the same as that configured for CC-V to allow both far-end and near-end defect conditions being resolved in a timeframe that has the same order of magnitude. This timeframe is application specific as described in Section 5.1.3. Methods of the out-of-band return paths will dictate how out-of-band RDIs are transmitted.

5.3. Alarm Reporting

The Alarm Reporting function, as required in Section 2.2.8 of RFC 5860 [11], relies upon an Alarm Indication Signal (AIS) packet to suppress alarms following detection of defect conditions at the server (sub-)layer.

When a server MEP asserts a signal fail condition, it notifies that to the co-located MPLS-TP client/server adaptation function that then generates OAM packets with AIS information in the downstream direction to allow the suppression of secondary alarms at the MPLS-TP MEP in the client (sub-)layer.

The generation of packets with AIS information starts immediately when the server MEP asserts a signal fail condition. These periodic OAM packets, with AIS information, continue to be transmitted until the signal fail condition is cleared.

It is assumed that to avoid spurious alarm generation a MEP detecting a loss of continuity defect (see Section 5.1.1.1) will wait for a hold-off interval prior to asserting an alarm to the management system. Therefore, upon receiving an OAM packet with AIS information, an MPLS-TP MEP enters an AIS defect condition and suppresses reporting of alarms to the NMS on the loss of continuity with its peer MEP, but it does not block traffic received from the transport path. A MEP resumes loss of continuity alarm generation upon detecting loss of continuity defect conditions in the absence of AIS condition.

MIPs, as well as intermediate nodes, do not process AIS information and forward these AIS OAM packets as regular data packets.

For example, let's consider a fiber cut between T-PE 1 and LSR 2 in the reference network of Figure 5. Assuming that all of the MEGs described in Figure 5 have proactive CC-V enabled, a LOC defect is detected by the MEPs of Sec12 SMEG, LSP13 LMEG, PW1 PSMEG, and PW1Z PMEG; however, in a transport network, only the alarm associated to the fiber cut needs to be reported to an NMS, while all secondary alarms should be suppressed (i.e., not reported to the NMS or reported as secondary alarms).

If the fiber cut is detected by the MEP in the physical layer (in LSR 2), LSR 2 can generate the proper alarm in the physical layer and suppress the secondary alarm associated with the LOC defect detected on Sec12 SMEG. As both MEPs reside within the same node, this process does not involve any external protocol exchange. Otherwise,

if the physical layer does not have enough OAM capabilities to detect the fiber cut, the MEP of Sec12 SMEG in LSR 2 will report a LOC alarm.

In both cases, the MEP of Sec12 SMEG in LSR 2 notifies the adaptation function for LSP13 LMEG that then generates AIS packets on the LSP13 LMEG in order to allow its MEP in S-PE 3 to suppress the LOC alarm. S-PE 3 can also suppress the secondary alarm on PW13 PSMEG because the MEP of PW13 PSMEG resides within the same node as the MEP of LSP13 LMEG. The MEP of PW13 PSMEG in S-PE 3 also notifies the adaptation function for PW1Z PMEG that then generates AIS packets on PW1Z PMEG in order to allow its MEP in T-PE Z to suppress the LOC alarm.

The generation of AIS packets for each MEG in the MPLS-TP client (sub-)layer is configurable (i.e., the operator can enable/disable the AIS generation).

The AIS condition is cleared if no AIS packet has been received in 3.5 times the AIS transmission period.

The AIS transmission period is traditionally one per second, but an option to configure longer periods would be also desirable. As a consequence, OAM packets need to self-identify the transmission period such that proper exit criteria can be established.

AIS packets are transmitted with the "minimum loss probability PHB" within a single network operator. For E-LSPs, this PHB is configurable on network operator's basis, while for L-LSPs, this is determined as per RFC 3270 [23].

5.4. Lock Reporting

The Lock Reporting function, as required in Section 2.2.7 of RFC 5860 [11], relies upon a Lock Report (LKR) packet used to suppress alarms following administrative locking action in the server (sub-)layer.

When a server MEP is locked, the MPLS-TP client (sub-)layer adaptation function generates packets with LKR information to allow the suppression of secondary alarms at the MEPs in the client (sub-)layer. Again, it is assumed that there is a hold-off for any loss of continuity alarms in the client-layer MEPs downstream of the node originating the Lock Report. In case of client (sub-)layer co-routed bidirectional transport paths, the LKR information is sent on both directions. In case of client (sub-)layer unidirectional transport paths, the LKR information is sent only in the downstream direction. As a consequence, in case of client (sub-)layer point-to-multipoint transport paths, the LKR information is sent only to the

MEPs that are downstream from the server (sub-)layer that has been administratively locked. Client (sub-)layer associated bidirectional transport paths behave like co-routed bidirectional transport paths if the server (sub-)layer that has been administratively locked is used by both directions; otherwise, they behave like unidirectional transport paths.

The generation of packets with LKR information starts immediately when the server MEP is locked. These periodic packets, with LKR information, continue to be transmitted until the locked condition is cleared.

Upon receiving a packet with LKR information, an MPLS-TP MEP enters an LKR defect condition and suppresses the loss of continuity alarm associated with its peer MEP but does not block traffic received from the transport path. A MEP resumes loss of continuity alarm generation upon detecting loss of continuity defect conditions in the absence of the LKR condition.

MIPs, as well as intermediate nodes, do not process the LKR information; they forward these LKR OAM packets as regular data packets.

For example, let's consider the case where the MPLS-TP Section between T-PE 1 and LSR 2 in the reference network of Figure 5 is administratively locked at LSR 2 (in both directions).

Assuming that all the MEGs described in Figure 5 have proactive CC-V enabled, a LOC defect is detected by the MEPs of LSP13 LMEG, PW1 PSMEG, and PW1Z PMEG; however, in a transport network all these secondary alarms should be suppressed (i.e., not reported to the NMS or reported as secondary alarms).

The MEP of Sec12 SMEG in LSR 2 notifies the adaptation function for LSP13 LMEG that then generates LKR packets on the LSP13 LMEG in order to allow its MEPs in T-PE 1 and S-PE 3 to suppress the LOC alarm. S-PE 3 can also suppress the secondary alarm on PW13 PSMEG because the MEP of PW13 PSMEG resides within the same node as the MEP of LSP13 LMEG. The MEP of PW13 PSMEG in S-PE 3 also notifies the adaptation function for PW1Z PMEG that then generates AIS packets on PW1Z PMEG in order to allow its MEP in T-PE Z to suppress the LOC alarm.

The generation of LKR packets for each MEG in the MPLS-TP client (sub-)layer is configurable (i.e., the operator can enable/disable the LKR generation).

The locked condition is cleared if no LKR packet has been received for 3.5 times the transmission period.

The LKR transmission period is traditionally one per second, but an option to configure longer periods would be also desirable. As a consequence, OAM packets need to self-identify the transmission period such that proper exit criteria can be established.

LKR packets are transmitted with the "minimum loss probability PHB" within a single network operator. For E-LSPs, this PHB is configurable on network operator's basis, while for L-LSPs, this is determined as per RFC 3270 [23].

5.5. Packet Loss Measurement

Packet Loss Measurement (LM) is one of the capabilities supported by the MPLS-TP Performance Monitoring (PM) function in order to facilitate reporting of Quality of Service (QoS) information for a transport path as required in Section 2.2.11 of RFC 5860 [11]. LM is used to exchange counter values for the number of ingress and egress packets transmitted and received by the transport path monitored by a pair of MEPs.

Proactive LM is performed by periodically sending LM OAM packets from a MEP to a peer MEP and by receiving LM OAM packets from the peer MEP (if a co-routed or associated bidirectional transport path) during the lifetime of the transport path. Each MEP performs measurements of its transmitted and received user data packets. These measurements are then correlated in real time with the peer MEP in the ME to derive the impact of packet loss on a number of performance metrics for the ME in the MEG. The LM transactions are issued such that the OAM packets will experience the same PHB scheduling class as the measured traffic while transiting between the MEPs in the ME.

For a MEP, near-end packet loss refers to packet loss associated with incoming data packets (from the far-end MEP), while far-end packet loss refers to packet loss associated with egress data packets (towards the far-end MEP).

Proactive LM can be operated in two ways:

- One-way: a MEP sends an LM OAM packet to its peer MEP containing all the required information to facilitate near-end packet loss measurements at the peer MEP.

- Two-way: a MEP sends an LM OAM packet with an LM request to its peer MEP, which replies with an LM OAM packet as an LM response. The request/response LM OAM packets contain all the required

information to facilitate both near-end and far-end packet loss measurements from the viewpoint of the originating MEP.

One-way LM is applicable to both unidirectional and bidirectional (co-routed or associated) transport paths, while two-way LM is applicable only to bidirectional (co-routed or associated) transport paths.

MIPs, as well as intermediate nodes, do not process the LM information; they forward these proactive LM OAM packets as regular data packets.

5.5.1. Configuration Considerations

In order to support proactive LM, the transmission rate and, for E-LSPs, the PHB class (associated with the LM OAM packets originating from a MEP) need to be configured as part of the LM provisioning. LM OAM packets should be transmitted with the PHB that yields the lowest drop precedence within the measured PHB Scheduling Class (see RFC 3260 [17]), in order to maximize reliability of measurement within the traffic class.

If that PHB class is not an ordered aggregate where the ordering constraint is all packets with the PHB class being delivered in order, LM can produce inconsistent results.

Performance monitoring (e.g., LM) is only relevant when the transport path is defect free. CC-V contributes to the accuracy of PM statistics by permitting the defect-free periods to be properly distinguished. Therefore, support of proactive LM has implications on the CC-V transmission period (see Section 5.1.3).

5.5.2. Sampling Skew

If an implementation makes use of a hardware forwarding path that operates in parallel with an OAM processing path, whether hardware or software based, the packet and byte counts may be skewed if one or more packets can be processed before the OAM processing samples counters. If OAM is implemented in software, this error can be quite large.

5.5.3. Multilink Issues

If multilink is used at the ingress or egress of a transport path, there may not be a single packet-processing engine where an LM packet can be injected or extracted as an atomic operation while having accurate packet and byte counts associated with the packet.

In the case where multilink is encountered along the route of the transport path, the reordering of packets within the transport path can cause inaccurate LM results.

5.6. Packet Delay Measurement

Packet Delay Measurement (DM) is one of the capabilities supported by the MPLS-TP PM function in order to facilitate reporting of QoS information for a transport path as required in Section 2.2.12 of RFC 5860 [11]. Specifically, proactive DM is used to measure the long-term packet delay and packet delay variation in the transport path monitored by a pair of MEPs.

Proactive DM is performed by sending periodic DM OAM packets from a MEP to a peer MEP and by receiving DM OAM packets from the peer MEP (if a co-routed or associated bidirectional transport path) during a configurable time interval.

Proactive DM can be operated in two ways:

- One-way: a MEP sends a DM OAM packet to its peer MEP containing all the required information to facilitate one-way packet delay and/or one-way packet delay variation measurements at the peer MEP. Note that this requires precise time synchronization at either MEP by means outside the scope of this framework.

- Two-way: a MEP sends a DM OAM packet with a DM request to its peer MEP, which replies with a DM OAM packet as a DM response. The request/response DM OAM packets contain all the required information to facilitate two-way packet delay and/or two-way packet delay variation measurements from the viewpoint of the originating MEP.

One-way DM is applicable to both unidirectional and bidirectional (co-routed or associated) transport paths, while two-way DM is applicable only to bidirectional (co-routed or associated) transport paths.

MIPs, as well as intermediate nodes, do not process the DM information; they forward these proactive DM OAM packets as regular data packets.

5.6.1. Configuration Considerations

In order to support proactive DM, the transmission rate and, for E-LSPs, the PHB (associated with the DM OAM packets originating from a MEP) need to be configured as part of the DM provisioning. DM OAM packets should be transmitted with the PHB that yields the lowest

drop precedence within the measured PHB Scheduling Class (see RFC 3260 [17]).

Performance monitoring (e.g., DM) is only relevant when the transport path is defect free. CC-V contributes to the accuracy of PM statistics by permitting the defect-free periods to be properly distinguished. Therefore, support of proactive DM has implications on the CC-V transmission period (see Section 5.1.3).

5.7. Client Failure Indication

The Client Failure Indication (CFI) function, as required in Section 2.2.10 of RFC 5860 [11], is used to help process client defects and propagate a client signal defect condition from the process associated with the local attachment circuit where the defect was detected (typically the source adaptation function for the local client interface). It is propagated to the process associated with the far-end attachment circuit (typically the source adaptation function for the far-end client interface) for the same transmission path, in case the client of the transport path does not support a native defect/alarm indication mechanism, e.g., AIS.

A source MEP starts transmitting a CFI to its peer MEP when it receives a local client signal defect notification via its local client signal fail indication. Mechanisms to detect local client signal fail defects are technology specific. Similarly, mechanisms to determine when to cease originating client signal fail indication are also technology specific.

A sink MEP that has received a CFI reports this condition to its associated client process via its local CFI function. Consequent actions toward the client attachment circuit are technology specific.

There needs to be a 1:1 correspondence between the client and the MEG; otherwise, when multiple clients are multiplexed over a transport path, the CFI packet requires additional information to permit the client instance to be identified.

MIPs, as well as intermediate nodes, do not process the CFI information; they forward these proactive CFI OAM packets as regular data packets.

5.7.1. Configuration Considerations

In order to support CFI indication, the CFI transmission rate and, for E-LSPs, the PHB of the CFI OAM packets should be configured as part of the CFI configuration.

6. OAM Functions for On-Demand Monitoring

In contrast to proactive monitoring, on-demand monitoring is initiated manually and for a limited amount of time, usually for operations such as diagnostics to investigate a defect condition.

On-demand monitoring covers a combination of "in-service" and "out-of-service" monitoring functions. The control and measurement implications are:

1. A MEG can be directed to perform an "on-demand" functions at arbitrary times in the lifetime of a transport path.

2. "Out-of-service" monitoring functions may require a priori configuration of both MEPs and intermediate nodes in the MEG (e.g., data-plane loopback) and the issuance of notifications into client layers of the transport path being removed from service (e.g., lock reporting)

3. The measurements resulting from "on-demand" monitoring are typically harvested in real time, as they are frequently initiated manually. These do not necessarily require different harvesting mechanisms than for harvesting proactive monitoring telemetry.

The functions that are exclusively out-of-service are those described in Section 6.3. The remainder are applicable to both in-service and out-of-service transport paths.

6.1. Connectivity Verification

The on-demand connectivity verification function, as required in Section 2.2.3 of RFC 5860 [11], is a transaction that flows from the originating MEP to a target MIP or MEP to verify the connectivity between these points.

Use of on-demand CV is dependent on the existence of a bidirectional ME or an associated return ME, or the availability of an out-of-band return path, because it requires the ability for target MIPs and MEPs to direct responses to the originating MEPs.

One possible use of on-demand CV would be to perform fault management without using proactive CC-V, in order to preserve network resources, e.g., bandwidth, processing time at switches. In this case, network management periodically invokes on-demand CV.

An additional use of on-demand CV would be to detect and locate a problem of connectivity when a problem is suspected or known to be based on other tools. In this case, the functionality will be triggered by the network management in response to a status signal or alarm indication.

On-demand CV is based upon generation of on-demand CV packets that should uniquely identify the MEG that is being checked. The on-demand functionality may be used to check either an entire MEG (end-to-end) or between the originating MEP and a specific MIP. This functionality may not be available for associated bidirectional transport paths or unidirectional paths, as the MIP may not have a return path to the originating MEP for the on-demand CV transaction.

When on-demand CV is invoked, the originating MEP issues a sequence of on-demand CV packets that uniquely identifies the MEG being verified. The number of packets and their transmission rate should be pre-configured at the originating MEP to take into account normal packet-loss conditions. The source MEP should use the mechanisms defined in Sections 3.3 and 3.4 when sending an on-demand CV packet to a target MEP or target MIP, respectively. The target MEP/MIP shall return a reply on-demand CV packet for each packet received. If the expected number of on-demand CV reply packets is not received at the originating MEP, this is an indication that a connectivity problem may exist.

On-demand CV should have the ability to carry padding such that a variety of MTU sizes can be originated to verify the MTU transport capability of the transport path.

MIPs that are not targeted by on-demand CV packets, as well as intermediate nodes, do not process the CV information; they forward these on-demand CV OAM packets as regular data packets.

6.1.1. Configuration Considerations

For on-demand CV, the originating MEP should support the configuration of the number of packets to be transmitted/received in each sequence of transmissions and their packet size.

In addition, when the CV packet is used to check connectivity toward a target MIP, the number of hops to reach the target MIP should be configured.

For E-LSPs, the PHB of the on-demand CV packets should be configured as well. This permits the verification of correct operation of QoS queuing as well as connectivity.

6.2. Packet Loss Measurement

On-demand Packet Loss Measurement (LM) is one of the capabilities supported by the MPLS-TP Performance Monitoring function in order to facilitate the diagnosis of QoS performance for a transport path, as required in Section 2.2.11 of RFC 5860 [11].

On-demand LM is very similar to proactive LM described in Section 5.5. This section focuses on the differences between on-demand and proactive LM.

On-demand LM is performed by periodically sending LM OAM packets from a MEP to a peer MEP and by receiving LM OAM packets from the peer MEP (if a co-routed or associated bidirectional transport path) during a pre-defined monitoring period. Each MEP performs measurements of its transmitted and received user data packets. These measurements are then correlated to evaluate the packet-loss performance metrics of the transport path.

Use of packet loss measurement in an out-of-service transport path requires a traffic source such as a test device that can inject synthetic traffic.

6.2.1. Configuration Considerations

In order to support on-demand LM, the beginning and duration of the LM procedures, the transmission rate, and, for E-LSPs, the PHB class (associated with the LM OAM packets originating from a MEP) must be configured as part of the on-demand LM provisioning. LM OAM packets should be transmitted with the PHB that yields the lowest drop precedence as described in Section 5.5.1.

6.2.2. Sampling Skew

The same considerations described in Section 5.5.2 for the proactive LM are also applicable to on-demand LM implementations.

6.2.3. Multilink Issues

Multilink issues are as described in Section 5.5.3.

6.3. Diagnostic Tests

Diagnostic tests are tests performed on a MEG that has been taken out of service.

6.3.1. Throughput Estimation

Throughput estimation is an on-demand out-of-service function, as required in Section 2.2.5 of RFC 5860 [11], that allows verifying the bandwidth/throughput of an MPLS-TP transport path (LSP or PW) before it is put in service.

Throughput estimation is performed between MEPs and between a MEP and a MIP. It can be performed in one-way or two-way modes.

According to RFC 2544 [12], this test is performed by sending OAM test packets at increasing rates (up to the theoretical maximum), computing the percentage of OAM test packets received, and reporting the rate at which OAM test packets begin to drop. In general, this rate is dependent on the OAM test packet size.

When configured to perform such tests, a source MEP inserts OAM test packets with a specified packet size and transmission pattern at a rate to exercise the throughput.

The throughput test can create congestion within the network, thus impacting other transport paths. However, the test traffic should comply with the traffic profile of the transport path under test, so the impact of the test will not be worse than the impact caused by the customers, whose traffic would be sent over that transport path, sending the traffic at the maximum rate allowed by their traffic profiles. Therefore, throughput tests are not applicable to transport paths that do not have a defined traffic profile, such as LSPs in a context where statistical multiplexing is leveraged for network capacity dimensioning.

For a one-way test, the remote sink MEP receives the OAM test packets and calculates the packet loss. For a two-way test, the remote MEP loops the OAM test packets back to the original MEP, and the local sink MEP calculates the packet loss.

It is worth noting that two-way throughput estimation is only applicable to bidirectional (co-routed or associated) transport paths and can only evaluate the minimum of available throughput of the two directions. In order to estimate the throughput of each direction uniquely, two one-way throughput estimation sessions have to be set up. One-way throughput estimation requires coordination between the transmitting and receiving test devices as described in Section 6 of RFC 2544 [12].

It is also worth noting that if throughput estimation is performed on transport paths that transit oversubscribed links, the test may not produce comprehensive results if viewed in isolation because the

impact of the test on the surrounding traffic needs to also be considered. Moreover, the estimation will only reflect the bandwidth available at the moment when the measure is made.

MIPs that are not targeted by on-demand test OAM packets, as well as intermediate nodes, do not process the throughput test information; they forward these on-demand test OAM packets as regular data packets.

6.3.1.1. Configuration Considerations

Throughput estimation is an out-of-service tool. The diagnosed MEG should be put into a locked state before the diagnostic test is started.

A MEG can be put into a locked state either via an NMS action or using the Lock Instruct OAM tool as defined in Section 7.

At the transmitting MEP, provisioning is required for a test signal generator that is associated with the MEP. At a receiving MEP, provisioning is required for a test signal detector that is associated with the MEP.

In order to ensure accurate measurement, care needs to be taken to enable throughput estimation only if all the MEPs within the MEG can process OAM test packets at the same rate as the payload data rates (see Section 6.3.1.2).

6.3.1.2. Limited OAM Processing Rate

If an implementation is able to process payload at much higher data rates than OAM test packets, then accurate measurement of throughput using OAM test packets is not achievable. Whether OAM packets can be processed at the same rate as payload is implementation dependent.

6.3.1.3. Multilink Considerations

If multilink is used, then it may not be possible to perform throughput measurement, as the throughput test may not have a mechanism for utilizing more than one component link of the aggregated link.

6.3.2. Data-Plane Loopback

Data-plane loopback is an out-of-service function, as required in Section 2.2.5 of RFC 5860 [11]. This function consists in placing a transport path, at either an intermediate or terminating node, into a data-plane loopback state, such that all traffic (including both

payload and OAM) received on the looped back interface is sent on the reverse direction of the transport path. The traffic is looped back unmodified except for normal per-hop processing such as TTL decrement.

The data-plane loopback function requires that the MEG is locked such that user data traffic is prevented from entering/exiting that MEG. Instead, test traffic is inserted at the ingress of the MEG. This test traffic can be generated from an internal process residing within the ingress node or injected by external test equipment connected to the ingress node.

It is also normal to disable proactive monitoring of the path as the MEP located upstream with respect to the node set in the data-plane loopback mode will see all the OAM packets originated by itself, and this may interfere with other measurements.

The only way to send an OAM packet (e.g., to remove the data-plane loopback state) to the MIPs or MEPs hosted by a node set in the data-plane loopback mode is via TTL expiry. It should also be noted that MIPs can be addressed with more than one TTL value on a co-routed bidirectional path set into data-plane loopback.

If the loopback function is to be performed at an intermediate node, it is only applicable to co-routed bidirectional paths. If the loopback is to be performed end to end, it is applicable to both co-routed bidirectional and associated bidirectional paths.

It should be noted that data-plane loopback function itself is applied to data-plane loopback points that can reside on different interfaces from MIPs/MEPs. Where a node implements data-plane loopback capability and whether it implements it in more than one point is implementation dependent.

6.3.2.1. Configuration Considerations

Data-plane loopback is an out-of-service tool. The MEG that defines a diagnosed transport path should be put into a locked state before the diagnostic test is started. However, a means is required to permit the originated test traffic to be inserted at the ingress MEP when data-plane loopback is performed.

A transport path, at either an intermediate or terminating node, can be put into data-plane loopback state via an NMS action or using an OAM tool for data-plane loopback configuration.

If the data-plane loopback point is set somewhere at an intermediate point of a co-routed bidirectional transport path, the side of the loopback function (east/west side or both sides) needs to be configured.

6.4. Route Tracing

It is often necessary to trace a route covered by a MEG from an originating MEP to the peer MEP(s) including all the MIPs in between. This may be conducted after provisioning an MPLS-TP transport path for, e.g., troubleshooting purposes such as fault localization.

The route tracing function, as required in Section 2.2.4 of RFC 5860 [11], is providing this functionality. Based on the fate-sharing requirement of OAM flows, i.e., OAM packets receive the same forwarding treatment as data packets, route tracing is a basic means to perform connectivity verification and, to a much lesser degree, continuity check. For this function to work properly, a return path must be present.

Route tracing might be implemented in different ways, and this document does not preclude any of them.

Route tracing should always discover the full list of MIPs and of peer MEPs. In case a defect exists, the route tracing function will only be able to trace up to the defect, and it needs to be able to return the incomplete list of OAM entities that it was able to trace so that the fault can be localized.

6.4.1. Configuration Considerations

The configuration of the route tracing function must at least support the setting of the number of trace attempts before it gives up.

6.5. Packet Delay Measurement

Packet Delay Measurement (DM) is one of the capabilities supported by the MPLS-TP PM function in order to facilitate reporting of QoS information for a transport path, as required in Section 2.2.12 of RFC 5860 [11]. Specifically, on-demand DM is used to measure packet delay and packet delay variation in the transport path monitored by a pair of MEPs during a pre-defined monitoring period.

On-demand DM is performed by sending periodic DM OAM packets from a MEP to a peer MEP and by receiving DM OAM packets from the peer MEP (if a co-routed or associated bidirectional transport path) during a configurable time interval.

On-demand DM can be operated in two modes:

- One-way: a MEP sends a DM OAM packet to its peer MEP containing all the required information to facilitate one-way packet delay and/or one-way packet delay variation measurements at the peer MEP. Note that this requires precise time synchronization at either MEP by means outside the scope of this framework.

- Two-way: a MEP sends a DM OAM packet with a DM request to its peer MEP, which replies with a DM OAM packet as a DM response. The request/response DM OAM packets contain all the required information to facilitate two-way packet delay and/or two-way packet delay variation measurements from the viewpoint of the originating MEP.

MIPs, as well as intermediate nodes, do not process the DM information; they forward these on-demand DM OAM packets as regular data packets.

6.5.1. Configuration Considerations

In order to support on-demand DM, the beginning and duration of the DM procedures, the transmission rate and, for E-LSPs, the PHB (associated with the DM OAM packets originating from a MEP) need to be configured as part of the DM provisioning. DM OAM packets should be transmitted with the PHB that yields the lowest drop precedence within the measured PHB Scheduling Class (see RFC 3260 [17]).

In order to verify different performances between long and short packets (e.g., due to the processing time), it should be possible for the operator to configure the packet size of the on-demand OAM DM packet.

7. OAM Functions for Administration Control

7.1. Lock Instruct

The Lock Instruct (LKI) function, as required in Section 2.2.6 of RFC 5860 [11], is a command allowing a MEP to instruct the peer MEP(s) to put the MPLS-TP transport path into a locked condition.

This function allows single-side provisioning for administratively locking (and unlocking) an MPLS-TP transport path.

Note that it is also possible to administratively lock (and unlock) an MPLS-TP transport path using two-side provisioning, where the NMS administratively puts both MEPs into an administrative lock condition. In this case, the LKI function is not required/used.

MIPs, as well as intermediate nodes, do not process the Lock Instruct information; they forward these on-demand LKI OAM packets as regular data packets.

7.1.1. Locking a Transport Path

A MEP, upon receiving a single-side administrative lock command from an NMS, sends an LKI request OAM packet to its peer MEP(s). It also puts the MPLS-TP transport path into a locked state and notifies its client (sub-)layer adaptation function upon the locked condition.

A MEP, upon receiving an LKI request from its peer MEP, can either accept or reject the instruction and replies to the peer MEP with an LKI reply OAM packet indicating whether or not it has accepted the instruction. This requires either an in-band or out-of-band return path. The LKI reply is needed to allow the MEP to properly report to the NMS the actual result of the single-side administrative lock command.

If the lock instruction has been accepted, it also puts the MPLS-TP transport path into a locked state and notifies its client (sub-)layer adaptation function upon the locked condition.

Note that if the client (sub-)layer is also MPLS-TP, Lock Report (LKR) generation at the client MPLS-TP (sub-)layer is started, as described in Section 5.4.

7.1.2. Unlocking a Transport Path

A MEP, upon receiving a single-side administrative unlock command from NMS, sends an LKI removal request OAM packet to its peer MEP(s).

The peer MEP, upon receiving an LKI removal request, can either accept or reject the removal instruction and replies with an LK removal reply OAM packet indicating whether or not it has accepted the instruction. The LKI removal reply is needed to allow the MEP to properly report to the NMS the actual result of the single-side administrative unlock command.

If the lock removal instruction has been accepted, it also clears the locked condition on the MPLS-TP transport path and notifies its client (sub-)layer adaptation function of this event.

The MEP that has initiated the LKI clear procedure, upon receiving a positive LKI removal reply, also clears the locked condition on the MPLS-TP transport path and notifies this event to its client (sub-)layer adaptation function.

Note that if the client (sub-)layer is also MPLS-TP, Lock Report (LKR) generation at the client MPLS-TP (sub-)layer is terminated, as described in Section 5.4.

8. Security Considerations

A number of security considerations are important in the context of OAM applications.

OAM traffic can reveal sensitive information, such as performance data and details, about the current state of the network. Insertion or modification of OAM transactions can mask the true operational state of the network, and in the case of transactions for administration control, such as lock or data-plane loopback instructions, these can be used for explicit denial-of-service attacks. The effect of such attacks is mitigated only by the fact that, for in-band messaging, the managed entities whose state can be masked is limited to those that transit the point of malicious access to the network internals due to the fate-sharing nature of OAM messaging. This is not true when an out-of-band return path is employed.

The sensitivity of OAM data therefore suggests that one solution is that some form of authentication, authorization, and encryption is in place. This will prevent unauthorized access to vital equipment, and it will prevent third parties from learning about sensitive information about the transport network. However, it should be observed that the combination of the frequency of some OAM transactions, the need for timeliness of OAM transaction exchange, and all permutations of unique MEP to MEP, MEP to MIP, and intermediate-system-originated transactions mitigates against the practical establishment and maintenance of a large number of security associations per MEG either in advance or as required.

For this reason, it is assumed that the internal links of the network are physically secured from malicious access such that OAM transactions scoped to fault and performance management of individual MEGs are not encumbered with additional security. Further, it is assumed in multi-provider cases where OAM transactions originate outside of an individual provider's trusted domain that filtering mechanisms or further encapsulation will need to constrain the potential impact of malicious transactions. Mechanisms that the framework does not specify might be subject to additional security considerations.

In case of misconfiguration, some nodes can receive OAM packets that they cannot recognize. In such a case, these OAM packets should be silently discarded in order to avoid malfunctions whose effects may

be similar to malicious attacks (e.g., degraded performance or even failure). Further considerations about data-plane attacks via G-ACh are provided in RFC 5921 [8].

9. Acknowledgments

The authors would like to thank all members of the teams (the Joint Working Team, the MPLS Interoperability Design Team in IETF, and the Ad Hoc Group on MPLS-TP in ITU-T) involved in the definition and specification of the MPLS Transport Profile.

The editors gratefully acknowledge the contributions of Adrian Farrel, Yoshinori Koike, Luca Martini, Yuji Tochio, and Manuel Paul for the definition of per-interface MIPs and MEPs.

The editors gratefully acknowledge the contributions of Malcolm Betts, Yoshinori Koike, Xiao Min, and Maarten Vissers for the Lock Report and Lock Instruct descriptions.

The authors would also like to thank Alessandro D'Alessandro, Loa Andersson, Malcolm Betts, Dave Black, Stewart Bryant, Rui Costa, Xuehui Dai, John Drake, Adrian Farrel, Dan Frost, Xia Liang, Liu Gouman, Peng He, Russ Housley, Feng Huang, Su Hui, Yoshionori Koike, Thomas Morin, George Swallow, Yuji Tochio, Curtis Villamizar, Maarten Vissers, and Xuequin Wei for their comments and enhancements to the text.

10. References

10.1. Normative References

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

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

[3] Nadeau, T., Ed., and C. Pignataro, Ed., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.

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

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

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

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

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

[9] Bocci, M., Levrau, L., and D. Frost, "MPLS Transport Profile User-to-Network and Network-to-Network Interfaces", RFC 6215, April 2011.

[10] Frost, D., Ed., Bryant, S., Ed., and M. Bocci, Ed., "MPLS Transport Profile Data Plane Architecture", RFC 5960, August 2010.

[11] Vigoureux, M., Ed., Ward, D., Ed., and M. Betts, Ed., "Requirements for Operations, Administration, and Maintenance (OAM) in MPLS Transport Networks", RFC 5860, May 2010.

[12] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, March 1999.

[13] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Service", RFC 2475, December 1998.

[14] ITU-T Recommendation G.806 (01/09), "Characteristics of transport equipment - Description methodology and generic functionality", January 2009.

10.2. Informative References

[15] Sprecher, N. and L. Fang, "An Overview of the OAM Tool Set for MPLS based Transport Networks", Work in Progress, June 2011.

[16] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.

[17] Grossman, D., "New Terminology and Clarifications for Diffserv", RFC 3260, April 2002.

[18] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering (TE)", RFC 4201, October 2005.

[19] ITU-T Recommendation G.707/Y.1322 (01/07), "Network node interface for the synchronous digital hierarchy (SDH)", January 2007.

[20] ITU-T Recommendation G.805 (03/00), "Generic functional architecture of transport networks", March 2000.

[21] ITU-T Recommendation Y.1731 (02/08), "OAM functions and mechanisms for Ethernet based networks", February 2008.

[22] IEEE Standard 802.1AX-2008, "IEEE Standard for Local and Metropolitan Area Networks - Link Aggregation", November 2008.

[23] 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.

[24] Bocci, M., Swallow, G., and E. Gray, "MPLS Transport Profile (MPLS-TP) Identifiers", RFC 6370, September 2011.

[25] Winter, R., Ed., van Helvoort, H., and M. Betts, "MPLS-TP Identifiers Following ITU-T Conventions", Work in Progress, July 2011.

11. Contributing Authors

Ben Niven-Jenkins
Velocix

EMail: ben@niven-jenkins.co.uk

Annamaria Fulignoli
Ericsson

EMail: annamaria.fulignoli@ericsson.com

Enrique Hernandez-Valencia
Alcatel-Lucent

EMail: Enrique.Hernandez@alcatel-lucent.com

Lieven Levrau
Alcatel-Lucent

EMail: Lieven.Levrau@alcatel-lucent.com

Vincenzo Sestito
Alcatel-Lucent

EMail: Vincenzo.Sestito@alcatel-lucent.com

Nurit Sprecher
Nokia Siemens Networks

EMail: nurit.sprecher@nsn.com

Huub van Helvoort
Huawei Technologies

EMail: hhelvoort@huawei.com

Martin Vigoureux
Alcatel-Lucent

EMail: Martin.Vigoureux@alcatel-lucent.com

Yaacov Weingarten
Nokia Siemens Networks

EMail: yaacov.weingarten@nsn.com

Rolf Winter
NEC

EMail: Rolf.Winter@nw.neclab.eu

Authors' Addresses

Dave Allan
Ericsson

EMail: david.i.allan@ericsson.com

Italo Busi
Alcatel-Lucent

EMail: Italo.Busi@alcatel-lucent.com