MPLS Working Group I. Busi (Ed)
Internet Draft Alcatel-Lucent
Intended status: Informational B. Niven-Jenkins (Ed)
BT
D. Allan (Ed)
Ericsson

Expires: April 2010 October 26, 2009

MPLS-TP OAM Framework
draft-ietf-mpls-tp-oam-framework-02.txt

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

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

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress".

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

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

Abstract

Multi-Protocol Label Switching (MPLS) Transport Profile (MPLS-TP) is based on a profile of the MPLS and pseudowire (PW) procedures as specified in the MPLS Traffic Engineering (MPLS-TE), pseudowire (PW) and multi-segment PW (MS-PW) architectures complemented with additional Operations, Administration and Maintenance (OAM) procedures for fault, performance and protection-switching management for packet transport applications that do not rely on the presence of a control plane.

This document describes a framework to support a comprehensive set of OAM procedures that fulfills the MPLS-TP OAM requirements [12].

Table of Contents

1. Introduction

As noted in [8], MPLS-TP defines a profile of the MPLS-TE and (MS-)PW architectures defined in RFC 3031 [2], RFC 3985 [5] and [7] which is complemented with additional OAM mechanisms and procedures for alarm, fault, performance and protection-switching management for packet transport applications.

[Editor's note - The draft needs to be reviewed to ensure support of OAM for p2mp transport paths]

In line with [13], 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.

The MPLS-TP OAM framework defined in this document provides a comprehensive set of OAM procedures that satisfy the MPLS-TP OAM requirements [12]. In this regard, it defines similar OAM functionality as for existing SONET/SDH and OTN OAM mechanisms (e.g. [16]).

1.1. Contributing Authors

Dave Allan, Italo Busi, Ben Niven-Jenkins, Annamaria Fulignoli, Enrique Hernandez-Valencia, Lieven Levrau, Dinesh Mohan, Vincenzo Sestito, Nurit Sprecher, Huub van Helvoort, Martin Vigoureux, Yaacov Weingarten, Rolf Winter

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [1].

2.1. Terminology

AC Attachment Circuit

DBN Domain Border Node

FDI Forward Defect Indication

LER Label Edge Router

LME LSP Maintenance Entity

LSP Label Switched Path

LSR Label Switch Router

LTCME LSP Tandem Connection Maintenance Entity

[Editor's note - Difference or similarity between tandem connection monitoring (TCM)_and Path Segment Tunnel (PST) need to be defined and agreed]

ME Maintenance Entity

MEG Maintenance Entity Group

MEP Maintenance Entity Group End Point

MIP Maintenance Entity Group Intermediate Point

PHB Per-hop Behavior

PME PW Maintenance Entity

PTCME PW Tandem Connection Maintenance Entity

PSN Packet Switched Network

PW   Pseudowire

SLA Service Level Agreement

SME Section Maintenance Entity

2.2. Definitions

Note - the definitions in this section are intended to be in line with ITU-T recommendation Y.1731 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.

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

Maintenance Entity (ME): Some portion of a transport path that requires management bounded by two points, 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 transport path in an OAM domain.

MEP: A MEG end point (MEP) is capable of initiating (MEP Source) and terminating (MEP Sink) OAM messages for fault management and performance monitoring. MEPs reside at the boundaries of an ME (details in section 3.2).

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

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

MIP: A MEG intermediate point (MIP) terminates and processes OAM messages and may generate OAM messages in reaction to received OAM

messages. It never generates unsolicited OAM messages itself. A MIP resides within an MEG between MEPs (details in section 3.2).

OAM domain: A domain, as defined in [11], whose entities are grouped for the purpose of keeping the OAM confined within that domain.

Note - within the rest of this document the term "domain" is used to indicate an "OAM domain"

OAM flow: Is the set of all OAM messages originating with a specific MEP that instrument one direction of a MEG.

OAM information element: An atomic piece of information exchanged between MEPs in MEG used by an OAM application.

OAM Message: One or more OAM information elements that when exchanged between MEPs or between MEPs and MIPs performs some OAM functionality (e.g. continuity check or connectivity verification)

OAM Packet: A packet that carries one or more OAM messages (i.e. OAM information elements).

Path: See Transport Path

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.

Tandem Connection: A tandem connection is an arbitrary part of a transport path that can be monitored (via OAM) independently from 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.

The following terms are defined in RFC 5654 [11] as follows:

Associated bidirectional path: A path that supports traffic flow in both directions but that is constructed from a pair of unidirectional paths (one for each direction) that are associated with one another at the path's ingress/egress points. The forward and backward directions are setup, monitored, and protected independently. As a consequence, they may or may not follow the same route (links and nodes) across the network.

Concatenated Segment: A serial-compound link connection as defined in G.805 [17]. A concatenated segment is a contiguous part of an LSP or

multi-segment PW that comprises a set of segments and their interconnecting nodes in sequence. See also "Segment".

Co-routed bidirectional path: A path where the forward and backward directions follow the same route (links and nodes) across the network. Both directions are setup, monitored and protected as a single entity. A transport network path is typically co-routed.

Layer network: Layer network is defined in G.805 [17]. A layer network provides for the transfer of client information and independent operation of the client OAM. A layer network may be described in a service context as follows: one layer network may provide a (transport) service to higher client layer network and may, in turn, be a client to a lower-layer network. A layer network is a logical construction somewhat independent of arrangement or composition of physical network elements. A particular physical network element may topologically belong to more than one layer network, depending on the actions it takes on the encapsulation associated with the logical layers (e.g., the label stack), and thus could be modeled as multiple logical elements. A layer network may consist of one or more sublayers. Section 1.4 (of RFC 5654) provides a more detailed overview of what constitutes a layer network. For additional explanation of how layer networks relate to the OSI concept of layering, see Appendix I of Y.2611 [19].

Section Layer Network: A section layer is a server layer (which may be MPLS-TP or a different technology) that provides for the transfer of the section-layer client information between adjacent nodes in the transport-path layer or transport service layer. A section layer may provide for aggregation of multiple MPLS-TP clients. Note that G.805 [17] defines the section layer as one of the two layer networks in a transmission-media layer network. The other layer network is the physical-media layer network.

Path: See Transport Path

Segment: A link connection as defined in G.805 [17]. A segment is the part of an LSP that traverses a single link or the part of a PW that traverses a single link (i.e., that connects a pair of adjacent {Switching|Terminating} Provider Edges). See also "Concatenated Segment". [editors: concept should be layer specific.. suggesting that the part of a PW that traverses a single physical link is a segment means a segment is pretty much bounded by duct ends, and by devices completely clueless as to the existence of the PW, visibility of the wrong layer, To group: we have a definition conflict between G.805 and usage of segment in IETF (e.g. PWE3), not sure how to resolve this, for discussion Nov 3]

Sublayer: Sublayer is defined in G.805 [17]. The distinction between a layer network and a sublayer is that a sublayer is not directly accessible to clients outside of its encapsulating layer network and offers no direct transport service for a higher layer (client) network. [editors: messy definition as it is context specific. Given MPLS has no PID, the transport path will always exist in a sublayer as the PW or PID label which has no forwarding context will be bottom of stack. Whether or not you actually think of the PW label as being a sublayer itself entirely dependant on usage SS or MS-PW, for discussion Nov 3rd]

Transport Path: A network connection as defined in G.805 [17]. In an MPLS-TP environment, a transport path corresponds to an LSP or a PW.

Transport Path Layer: A (sub)layer network that provides point-to-point or point-to-multipoint transport paths. It is instrumented with OAM mechanisms that are independent of the clients it is transporting. [editor: if you look at the sublayer discussion above, this term pretty much universally must be a transport path sub-layer. The transport path cannot be a layer to itself in the MPLS_TP architecture unless we are discussing multi-segment dry martini, for discussion Nov 3rd]

Unidirectional path: A path that supports traffic flow in only one direction.

The term 'Per-hop Behavior' is defined in [14] as follows:

Per-hop Behavior: a description of the externally observable forwarding treatment applied at a differentiated services-compliant node to a behavior aggregate.

3. Functional Components

MPLS-TP defines a profile of the MPLS and PW architectures ([2], [5] and [7]) that is designed to transport service traffic where the characteristics of information transfer between the transport path endpoints can be demonstrated to comply with certain performance and quality guarantees. In order to verify and maintain these performance and quality guarantees, there is a need to not only apply OAM functionality 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 path.

In order to describe the required OAM functionality, this document introduces a set of high-level functional components. [Note - discussion in Munich -tues concluded that TCM not possible with PWs -

can monitor a single PW segment - but attempting to monitor more than one segment converts the PW into an LSP and therefore the intervening SPEs are unable to see the PW as a PW due to the differences in how OAM flows are disambiguated.] [editors: if true this IMO is a huge problem as the one place I would really want TCM is a multi-domain MS-PW, else I have to control plane peer at two layers, for discussion Nov 3rd]

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

These functional components should be instantiated when the path is created by either the management plane or by the control plane (if present). Some components may be instantiated after the path is initially created (e.g. TCM).

3.1. Maintenance Entity (ME) and Maintenance Entity Group (MEG)

[editors: rather than fight chicken and egg, we made two sections into one]

MPLS-TP OAM operates in the context of Maintenance Entities (MEs) that are a relationship between two points of a point to point [editors: why has this restriction been added, for discussion Nov 3rd] transport path to which maintenance and monitoring operations apply. These two points are called Maintenance Entity Group End Points (MEPs). In between these two points zero or more intermediate points, called Maintenance Entity Group MEG Intermediate Points (MIPS), MAY exist and can be shared by more than one ME in a MEG.

The MEPs that form an MEG are configured and managed to limit the scope of an OAM flow within the MEG the MEPs belong to (i.e. within the domain of the transport path or segment, in the specific layer network, that is being monitored and managed). A misbranching fault may cause OAM packets to be delivered to a MEP that is not in the MEG of origin.

The abstract reference model for an ME with MEPs and MIPs is described in Figure 1 below:

```
+-+    +-+    +-+    +-+
|A|----|B|----|C|----|D|
+-+    +-+    +-+    +-+
```

Figure 1 ME Abstract Reference Model

The instantiation of this abstract model to different MPLS-TP entities is described in section 4. In this model, nodes A, B, C and D can be LER/LSR for an LSP or the {S|T}-PEs for a MS-PW. MEPs reside in nodes A and D while MIPs reside in nodes B and C. The links connecting adjacent nodes can be physical links, sub-layer LSPs or lower layer TCMs.

This functional model defines the relationships between all OAM entities from a maintenance perspective, to allow each Maintenance Entity to monitor and manage the layer network under its responsibility and to localize problems efficiently.

[Dave: given how these definitions are shaking out, should the MEG and ME not be confined to a sub-layer, there is no such thing as a completely self contained "layer" in the architecture to which a MEG can apply, for Nov 3rd]

An MPLS-TP maintenance entity group can cover either the whole end-to-end or a Tandem Connection of the transport path. A Maintenance Entity Group may be defined to monitor the transport path for fault and/or performance management.

In case of associated bi-directional paths, two independent 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 source MEP does not exist in a unidirectional ME.

The following properties apply to all MPLS-TP MEs:

o They can be nested but not overlapped, e.g. an ME may cover a segment or a concatenated segment of another ME, and may also include the forwarding engine(s) of the node(s) at the edge(s) of the segment or concatenated segment, but all its MEPs and MIPs are no longer part of the encompassing ME. It is possible that MEPs of nested MEs reside on a single node.

o Each OAM flow is associated with a single Maintenance Entity.

o OAM packets are subject to the same forwarding treatment (i.e. fate share) as the data traffic and in some cases may be required to have common queuing discipline E2E with the class of traffic monitored. OAM packets can be distinguished from the data traffic using the GAL and ACH constructs [9] for LSP and Section or the ACH construct [6]and [9] for (MS-)PW.

[Propose from Munich - rewrite to describe the MEG as collection of one or more maint entities and then immediately define an ME.

[editors: much of this comment is actually either ME or MEP/MIP specific, not MEG specific, hence we are struggling as to what to do with this, for discussion Nov 3rd]

A key point in the definition of an ME is the end-points are defined by location of the logical function MEP

Later in the framework we will discuss the precision with which we can identify the location of a MEP/MIP i.e, ingress i/f, egress i/f or node.

We need to distinguish between the point of interception of an OAM msg and the point where the action takes place.

Somewhere we need to distinguish between the OAM control function and the OAM measurement function. i.e. we set up a loop back (a control function, in which case the OAM message may be intercepted and actioned anywhere convenient), and the measurement function (i.e. looping the packet to determine that it reached a particular part of the network) which needs to be actioned at a precisely know and stipulated point in the network/equipment.

Note that not all functionality / processing of an OAM pkt needs to take place at the point of measurement.

We considered that an OAM function can be decomposed into the following components

- Instruction or command

- Execution

- Addressing (node, interface etc) is ttl/LSP enough - do we need sub-addressing to cause execution on a specific component in the node - i.e. egress interface

- Response via OAM

- Reporting to mgt interface

It is useful to further decompose this into an initiator and a responder in general an initiator is the source mep and the responder is a mip or a sink mep. There are exceptions to this such as a mip initiating an AIS msg or lock indication.]

Another OAM construct is referred to as Maintenance Entity Group, which is a collection of one or more MEs that belongs to the same transport path and that are maintained and monitored as a group.

A use case for an MEG with more than one ME is point-to-multipoint OAM. The reference model for the p2mp MEG is represented in Figure 2.

```
                           +-+
                        /--|D|
                       /   +-+
                    +-+
                 /--|C|
      +-+    +-+/   +-+\   +-+
      |A|----|B|        \--|E|
      +-+    +-+\   +-+    +-+
                 \--|F|
                    +-+
```

Figure 2 Reference Model for p2mp MEG

In case of p2mp transport paths, the OAM operations are independent for each ME (A-D, A-E and A-F):

o Fault conditions - depending from where the failure is located

o Packet loss - depending from where the packets are lost

o Packet delay - depending on different paths

Each leaf (i.e. D, E and F) terminates OAM messages to monitor its own [Root, Leaf] ME while the root (i.e. A) generates OAM messages to monitor all the MEs of the p2mp MEG. In this particular case, the p2mp transport path is monitored by a MEG that consists of three MEs. Nodes B and C might implement a MIP in the corresponding MEGs.

3.2. MEG End Points (MEPs)

MEG End Points (MEPs) are the end points of an MEG. In the context of an MPLS-TP LSP, only LERs can implement MEPs while in the context of an LSP Tandem Connection both LERs and LSRs can implement MEPs. Regarding MPLS-TP PW, only T-PEs can implement MEPs while for a PW Tandem Connection both T-PEs and S-PEs can implement MEPs. In the context of MPLS-TP Section, any MPLS-TP NE can implement MEPs.

[Munich: See note about PW Tandem monitoring earlier, and whether a PW can be a tandem connection]

MEPs are responsible for activating and controlling all of the OAM functionality for the MEG. A MEP is capable of initiating and terminating OAM messages for fault management and performance monitoring. These OAM messages are encapsulated into an OAM packet using the G-ACh as defined in RFC 5586 [9]: in this case the G-ACh message is an OAM message and the channel type indicates an OAM message. A MEP terminates all the OAM packets it receives from the MEG it belongs to. The MPLS label identifies the MEG the OAM packet belongs to.

Once an MEG is configured, the operator can configure which OAM functions to use on the MEG but the MEPs are always enabled. A node at the edge of an MEG always support a MEP.

MEPs have to prevent OAM packets corresponding to a MEG from leaking outside that MEG:

o A MEP sink terminates all the OAM packets that it receives corresponding to its MEG and does not forward them further along the path.

o A MEP in a tandem connection tunnels all the OAM packets that it receives, upstream from the associated MEG to prevent them from being processed within the associated MEG. The usage of the label stacking mechanism allows all the MEPs and MIPs within the MEG to distinguish tunneled OAM packets from OAM packets that belong to that MEG.

MPLS-TP MEP passes a fault indication to its client (sub-)layer network as a consequent action of fault detection. [editor: interesting case, is this always sink, or are we considering loopbacks where inserting fault indication into the client (s)layer is comparatively useless. We wrestled with same problem with RSVP errors in the past ..., for Nov 3rd]

A MEP of an MPLS-TP transport path (Section, LSP or PW) 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 MEP source and MEP sink coincide with transport paths' source and sink terminations.

A MEP of an MPLS-TP tandem connection is not necessarily coincident with the termination of the MPLS-TP transport path (LSP or PW) and monitors the transport path for failures or performance degradation (e.g. based on packet counts) within the boundary of the MEG for the tandem connection.

It may occur in TCM that two MEPs are set on both sides of the forwarding engine such that the MEG is entirely internal to the node.

Note that a MEP can only exist at the beginning and end of a sub-layer i.e. an LSP or PW. If we need to add a monitoring point within an LSP we create a new sub-layer. We need to describe the migration process for adding a TCM segment.

We have the case of a MIP sending msg to a MEP. To do this it uses the LSP label - i.e. the top label of the stack at that point. [editors: clarify in section 3.4]

3.3. MEG Intermediate Points (MIPs)

A MEG Intermediate Point (MIP) is a point between the two MEPs of an ME.

A MIP is capable of reacting to some OAM packets and forwarding all the other OAM packets while ensuring fate sharing with data plane packets. However, a MIP does not initiate [unsolicited OAM - editors: this text was removed in the commented .rtf document from Munich but not tracked as a revision, validate this change Nov 3rd] packets, but may be addressed by OAM packets initiated by one of the MEPs of the ME. A MIP can generate OAM packets only in response to OAM packets that are sent on the MEG it belongs to.

An intermediate node within an MEG can either:

- o not support MPLS-TP OAM (i.e. no MIPs per node)

- o support per-node MIP (i.e. a single MIP per node)

- o support per-interface MIP (i.e. two MIPs per node on both sides of the forwarding engine)

A node at the edge of an MEG can also support a MEP and a per-interface MIP at the two sides of the forwarding engine.

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. It is always assumed that the "pipe" model of TTL handling is used by the MPLS transport profile.

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 inferred from the MPLS label.

Once an MEG is configured, the operator can enable/disable the MIPs on the nodes within the MEG.

3.4. Server MEPs

A server MEP is a MEP of an ME that is either:

- o defined in a layer network below the MPLS-TP layer network being referenced, or

- o defined in a sub-layer of the MPLS-TP layer network that is below the sub-layer being referenced.

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

A server MEP also interacts with the client/server adaptation function between the client (MPLS-TP) layer network and the server layer network. The adaptation function maintaints state on the mapping of MPLS-TP transport paths that are setup over that server layer's transport path.

For example, a server MEP can be either:

- o A termination point of a physical link (e.g. 802.3), an SDH VC or OTH ODU for the MPLS-TP Section layer network, defined in section 4.1;

- o An MPLS-TP Section MEP for MPLS-TP LSPs, defined in section 4.2;

- o An MPLS-TP LSP MEP for MPLS-TP PWs, defined in section 4.4;

- o An MPLS-TP LSP Tandem Connection MEP for higher-level LTCMEs, defined in section 4.3;

- o An MPLS-TP PW Tandem Connection MEP for higher-level PTCMEs, defined in section 4.5.

The server MEP can run appropriate OAM functions for fault detection within the server (sub-)layer network, and notifies a fault indication to its client MPLS-TP layer network. Server MEP OAM functions are outside the scope of this document.

3.5. Tandem Connection

A tandem connection is instantiated to support tandem connection monitoring (TCM).

TCM for a given portion of a transport path is implemented by first creating a hierarchical LSP that that has a 1:1 association with portion of the transport path that is to be uniquely monitored such that there is direct correlation between all FM and PM information gathered for the tandem connection AND the monitored portion of the E2E path. The tandem connection is monitored using normal LSP monitoring. There are a number of implications to this approach:

1) The hierarchical LSP would use the uniform model of EXP code point copying between sub-layers for diffserv such that the E2E markings and PHB treatment was preserved in the tandem connection.

2) The hierarchical LSP would use the pipe model for TTL handling such that MIP addressing for the E2E entity would be distinct from the tandem connection.

3) PM statistics need to be adjusted for the overhead of the additional sub-layer.

4) The server sub-layer LSP is viewed as single hop by the client LSP. The E2E ME source MEPs cannot direct transactions to tandem connection MIPs.

[editors: the text from Munich suggested that a tandem connection could be N:1, we've stuck with 1:1 such that there would be direct correlation of PM stats between the tandem connection and the monitored portion of the transport path, a N:1 hierarchical LSP IF WE INSIST on including, should be documented as a separate procedure]

4. Reference Model

The reference model for the MPLS-TP framework builds upon the concept of an MEG, and its associated MEPs and MIPs, to support the functional requirements specified in [12].

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

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

o A LSP Maintenance Entity (LME), allowing monitoring and management of an end-to-end LSP (between LERs).

o A PW Maintenance Entity (PME), allowing monitoring and management of an end-to-end SS/MS-PWs (between T-PEs).

o An LSP Tandem Connection Maintenance Entity (LTCME), allowing monitoring and management of an LSP Tandem Connection between any LER/LSR along the LSP. [Munich: Please clarify that an LTCME is JUST an ordinary hierarchical LSP (RFC3031).

Note - TCM only makes sense for LSPs as previously noted.]The MEs specified in this MPLS-TP framework are compliant with the architecture framework for MPLS MS-PWs [7] and MPLS LSPs [2].

Hierarchical LSPs are also supported. In this case, each LSP Tunnel in the hierarchy is a different sub-layer network that can be monitored independently from higher and lower level LSP tunnels in the hierarchy, end-to-end (from LER to LER) by an LME. Tandem Connection monitoring via LTCME are applicable on each LSP Tunnel in the hierarchy.

[Munich: There was discussion on above para - and it was suggested that it be removed.] [for discussion Nov 3rd, TCM and hierarchical LSPs...]

```
           Native  |<------------------- MS-PW1Z ------------------->|  Native
           Layer   |                                                 |   Layer
          Service  |    |<-PSN13->|    |<-PSN3X->|    |<-PSNXZ->|    |  Service
           (AC1)   V    V   LSP   V    V   LSP   V    V   LSP   V    V   (AC2)
                   +----+   +-+   +----+         +----+   +-+   +----+
     +----+        |TPE1|   | |   |SPE3|         |SPEX|   | |   |TPEZ|
+----+
     |    |        |    |=========|    |=========|    |=========|    |
|    |
     | CE1|--------|........PW13.......|...PW3X..|........PWXZ.......|-------|
CE2 |
     |    |        |    |=========|    |=========|    |=========|    |
|    |
     +----+        | 1  |   |2|   | 3  |         | X  |   |Y|   | Z  |
+----+

                   +----+   +-+   +----+         +----+   +-+   +----+
                   .                   .         .                   .
                   |                   |         |                   |
                   |<---- Domain 1 --->|         |<---- Domain Z --->|
                   ^------------------- PW1Z  PME -------------------^
                   ^---- PW13 PTCME ---^         ^---- PWXZ PTCME ---^
                        ^---------^                   ^---------^
                         PSN13 LME                     PSNXZ LME

                        ^---^ ^---^    ^---------^    ^---^ ^---^
                        Sec12 Sec23       Sec3X       SecXY SecYZ
                         SME   SME         SME         SME   SME

   TPE1: Terminating Provider Edge 1                 SPE2: Switching Provider
Edge 3
   TPEX: Terminating Provider Edge X                 SPEZ: Switching Provider
Edge Z

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

Figure 3 Reference Model for the MPLS-TP OAM Framework

Figure 3 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, LSR1 is adjacent to LSR2 via the MPLS Section Sec12 and LSR2 is adjacent to LSR3 via the MPLS Section Sec23. Similarly, in Domain Z, LSRX is adjacent to LSRY via the MPLS Section SecXY and LSRY is adjacent to LSRZ via the MPLS Section SecYZ. In addition, LSR3 is adjacent to LSRX via the MPLS Section 3X.

Figure 3 also shows a bi-directional MS-PW (PW1Z) between AC1 on TPE1 and AC2 on TPEZ. The MS-PW consists of three bi-directional PW Segments: 1) PW13 segment between T-PE1 and S-PE3 via the bi-directional PSN13 LSP, 2) PW3X segment between S-PE3 and S-PEX, via the bi-directional PSN3X LSP, and 3) PWXZ segment between S-PEX and T-PEZ via the bi-directional PSNXZ LSP.

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

Note that there are no constrains 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 3, it should be possible to configure one or more MEPs on the same node if that node is the endpoint of one or more MEGs.

Figure 3 does not describe a PW3X PTCME because typically TCMs are used to monitor an OAM domain (like PW13 and PWXZ PTCMEs)   rather than the segment between two OAM domains. However the OAM framework does not pose any constraints on the way TCM are instantiated as long as they are not overlapping.

The subsections below define the MEs specified in this MPLS-TP OAM architecture framework document. Unless otherwise stated, all references to domains, LSRs, MPLS Sections, LSPs, pseudowires and MEs in this section are made in relation to those shown in Figure 3.

4.1. MPLS-TP Section Monitoring

An MPLS-TP Section ME (SME) is an MPLS-TP maintenance entity intended to an MPLS Section as defined in [11]. An SME may be configured on any MPLS section. SME OAM packets must fate share with the user data packets sent over the monitored MPLS Section.

An SME 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 (and MPLS-TP enabled) LSRs rather than monitoring the individual LSP or PW segments traversing the MPLS Section and the server layer technology does not provide adequate OAM capabilities.

```
                   |<------------------- MS-PW1Z ------------------->|
                   |                                                 |
                   |    |<-PSN13->|    |<-PSN3X->|    |<-PSNXZ->|    |
                   V    V   LSP   V    V   LSP   V    V   LSP   V    V
                   +----+   +-+   +----+         +----+   +-+   +----+
     +----+        |TPE1|   | |   |SPE3|         |SPEX|   | |   |TPEZ|
+----+
     |    |  AC1   |    |=========|    |=========|    |=========|    |  AC2
|    |
     | CE1|--------|........PW13.......|...PW3X..|.......PWXZ........|-------|
CE2 |
     |    |        |    |=========|    |=========|    |=========|    |
|    |
     +----+        | 1  |   |2|   | 3  |         | X  |   |Y|   | Z  |
+----+
                   +----+   +-+   +----+         +----+   +-+   +----+

                        ^--^  ^--^     ^---------^     ^--^  ^--^
                        Sec12 Sec23       Sec3X       SecXY SecYZ
                         SME   SME         SME         SME   SME
```

Figure 4 Reference Example of MPLS-TP Section MEs (SME)

Figure 4 shows 5 Section MEs configured in the path between AC1 and AC2: 1) Sec12 ME associated with the MPLS Section between LSR 1 and LSR 2, 2) Sec23 ME associated with the MPLS Section between LSR 2 and LSR 3, 3) Sec3X ME associated with the MPLS Section between LSR 3 and LSR X, 4) SecXY ME associated with the MPLS Section between LSR X and LSR Y, and 5) SecYZ ME associated with the MPLS Section between LSR Y and LSR Z.

4.2. MPLS-TP LSP End-to-End Monitoring

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

An LME 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.

```
                   |<------------------- MS-PW1Z ------------------->|
                   |                                                 |
                   |    |<-PSN13->|    |<-PSN3X->|    |<-PSNXZ->|    |
                   V    V   LSP   V    V   LSP   V    V   LSP   V    V
                   +----+   +-+   +----+         +----+   +-+   +----+
     +----+        |TPE1|   | |   |SPE3|         |SPEX|   | |   |TPEZ|
+----+
     |    |  AC1   |    |=========|    |=========|    |=========|    |  AC2
|    |
     | CE1|--------|........PW13.......|...PW3X..|........PWXZ.......|-------|
CE2 |
     |    |        |    |=========|    |=========|    |=========|    |
|    |
     +----+        | 1  |   |2|   | 3  |         | X  |   |Y|   | Z  |
+----+
                   +----+   +-+   +----+         +----+   +-+   +----+

                        ^---------^                   ^---------^
                         PSN13 LME                     PSNXZ LME
```

Figure 5 Examples of MPLS-TP LSP MEs (LME)

Figure 5 depicts 2 LMEs configured in the path between AC1 and AC2: 1) the PSN13 LME between LER 1 and LER 3, and 2) the PSNXZ LME between LER X and LER Y. Note that the presence of a PSN3X LME in such a configuration is optional, hence, not precluded by this framework. For instance, the SPs may prefer to monitor the MPLS-TP Section between the two LSRs rather than the individual LSPs.

4.3. MPLS-TP LSP Tandem Connection Monitoring

An MPLS-TP LSP Tandem Connection Monitoring ME (LTCME) is an MPLS-TP maintenance entity intended to monitor an arbitrary part of an LSP between a given pair of LSRs independently from the end-to-end monitoring (LME). An LTCME can monitor an LSP segment or concatenated segment and it may also include the forwarding engine(s) of the node(s) at the edge(s) of the segment or concatenated segment.

Multiple LTCMEs MAY be configured on any LSP. The LSRs that terminate the LTCME may or may not be immediately adjacent at the MPLS-TP layer. LTCME OAM packets must fate share with the user data packets sent over the monitored LSP segment.

A LTCME can be defined between the following entities:

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

An LTCME is intended to be deployed in scenarios where it is preferable to monitor the behaviour of a part of an LSP 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.

Note that LTCMEs are equally applicable to hierarchical LSPs.

```
                   |<--------------------- PW1Z -------------------->|
                   |                                                 |
                   |    |<--------------PSN1Z LSP-------------->|    |
                   |    |<-PSN13->|    |<-PSN3X->|    |<-PSNXZ->|    |
                   V    V  S-LSP  V    V  S-LSP  V    V  S-LSP  V    V
                   +----+   +-+   +----+         +----+   +-+   +----+
     +----+        | PE1|   | |   |DBN3|         |DBNX|   | |   | PEZ|
+----+
     |    |  AC1   |    |=======================================|    |  AC2
|    |
     | CE1|--------|.......................PW1Z......................|-------|
CE2 |
     |    |        |    |=======================================|    |
|    |
     +----+        | 1  |   |2|   | 3  |         | X  |   |Y|   | Z  |
+----+

                   +----+   +-+   +----+         +----+   +-+   +----+
                   .                   .         .                   .
                   |                   |         |                   |
                   |<---- Domain 1 --->|         |<---- Domain Z --->|

                        ^---------^                   ^---------^
                        PSN13 LTCME                   PSNXZ LTCME
                        ^---------------------------------------^
                                        PSN1Z LME
```

DBN: Domain Border Node

Figure 6 MPLS-TP LSP Tandem Connection Monitoring ME (LTCME)

Figure 6 depicts a variation of the reference model in Figure 3 where there is an end-to-end PSN LSP (PSN1Z LSP) between PE1 and PEZ. PSN1Z LSP consists of, at least, three LSP Concatenated Segments: PSN13, PSN3X and PSNXZ. In this scenario there are two separate LTCMEs configured to monitor the PSN1Z LSP: 1) a LTCME monitoring the PSN13 LSP Concatenated Segment on Domain 1 (PSN13 LTCME), and 2) a LTCME monitoring the PSNXZ LSP Concatenated Segment on Domain Z (PSNXZ LTCME).

It is worth noticing that LTCMEs can coexist with the LME monitoring the end-to-end LSP and that LTCME MEPs and LME MEPs can be coincident in the same node (e.g. PE1 node supports both the PSN1Z LME MEP and the PSN13 LTCME MEP).

4.4. MPLS-TP PW Monitoring

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

A PME 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 aggregating multiple PWs between PEs.

```
                   |<------------------- MS-PW1Z ------------------->|
                   |                                                 |
                   |    |<-PSN13->|    |<-PSN3X->|    |<-PSNXZ->|    |
                   V    V   LSP   V    V   LSP   V    V   LSP   V    V
                   +----+   +-+   +----+         +----+   +-+   +----+
     +----+        |TPE1|   | |   |SPE3|         |SPEX|   | |   |TPEZ|
+----+
     |    |  AC1   |    |=========|    |=========|    |=========|    |  AC2
|    |
     | CE1|--------|........PW13.......|...PW3X..|........PWXZ.......|-------|
CE2 |
     |    |        |    |=========|    |=========|    |=========|    |
|    |
     +----+        | 1  |   |2|   | 3  |         | X  |   |Y|   | Z  |
+----+
                   +----+   +-+   +----+         +----+   +-+   +----+

                   ^--------------------PW1Z PME-------------------^
```

Figure 7 MPLS-TP PW ME (PME)

Figure 7 depicts a MS-PW (MS-PW1Z) consisting of three segments: PW13, PW3X and PWXZ and its associated end-to-end PME (PW1Z PME).

4.5. MPLS-TP MS-PW Tandem Connection Monitoring

An MPLS-TP MS-PW Tandem Connection Monitoring ME (PTCME) is an MPLS-TP maintenance entity intended to monitor an arbitrary part of an MS-PW between a given pair of PEs independently from the end-to-end monitoring (PME). A PTCME can monitor a PW segment or concatenated segment and it may also include the forwarding engine(s) of the node(s) at the edge(s) of the segment or concatenated segment.

Multiple PTCMEs MAY be configured on any MS-PW. The PEs may or may not be immediately adjacent at the MS-PW layer. PTCME OAM packets fate share with the user data packets sent over the monitored PW Segment.

A PTCME can be defined between the following entities:

o T-PE and any S-PE of a given MS-PW

o Any two S-PEs of a given MS-PW. It can span several PW segments.

A PTCME is intended to be deployed in scenarios where it is preferable to monitor the behaviour of a part of a MS-PW rather than the entire end-to-end PW itself, for example to monitor an MS-PW Segment within a given network domain of an inter-domain MS-PW.

```
                   |<------------------- MS-PW1Z ------------------->|
                   |                                                  |
                   |    |<-PSN13->|    |<-PSN3X->|    |<-PSNXZ->|    |
                   V    V   LSP   V    V   LSP   V    V   LSP   V    V
                   +----+   +-+   +----+         +----+   +-+   +----+
     +----+        |TPE1|   | |   |SPE3|         |SPEX|   | |   |TPEZ|
+----+
     |    |  AC1   |    |=========|    |=========|    |=========|    |  AC2
|    |
     | CE1|--------|........PW13.......|...PW3X..|........PWXZ.......|-------|
CE2 |
     |    |        |    |=========|    |=========|    |=========|    |
|    |
     +----+        | 1  |   |2|   | 3  |         | X  |   |Y|   | Z  |
+----+

                   +----+   +-+   +----+         +----+   +-+   +----+

                   ^---- PW1 PTCME ----^         ^---- PW5 PTCME ----^
                   ^---------------------PW1Z PME--------------------^
```

Figure 8 MPLS-TP MS-PW Tandem Connection Monitoring (PTCME)

Figure 8 depicts the same MS-PW (MS-PW1Z) between AC1 and AC2 as in Figure 7. In this scenario there are two separate PTCMEs configured to monitor MS-PW1Z: 1) a PTCME monitoring the PW13 MS-PW Segment on Domain 1 (PW13 PTCME), and 2) a PTCME monitoring the PWXZ MS-PW Segment on Domain Z with (PWXZ PTCME).

It is worth noticing that PTCMEs can coexist with the PME monitoring the end-to-end MS-PW and that PTCME MEPs and PME MEPs can be coincident in the same node (e.g. TPE1 node supports both the PW1Z PME MEP and the PW13 PTCME MEP).

5. OAM Functions for proactive monitoring

[Munich: Note the fwk needs to be explicit about the mapping of functions to the tools we have chosen.] [editors: shouldn't it be the other way around?]

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.

5.1. Continuity Check and Connectivity Verification

Proactive Continuity Check functions are used to detect a loss of continuity defect (LOC) between two MEPs in an MEG.

Proactive Connectivity Verification functions 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 of OAM packets by the source MEP that are processed by the sink MEP. As a consequence these two functions are grouped together into Continuity Check and Connectivity Verification (CC-V) OAM packets.

In order to perform pro-active Connectivity Verification function, each CC-V OAM packet MUST also include a globally unique Source MEP identifier. When used to perform only pro-active Continuity Check function, the CC-V OAM packet MAY not include any MEG identifier.

Different formats of MEP identifiers are defined in [10] to address different applications. When MPLS-TP is deployed in transport network applications as defined by ITU-T, the ICC-based format for MEP identification is the DEFAULT and MANDATORY identification scheme. When MPLS-TP is deployed in IP-based environment, the IP-based MEP identification is the DEFAULT and MANDATORY identification scheme.

As a consequence, it is not possible to detect misconnections between two MEGs monitored only for Continuity while it is possible to detect any misconnection between two MEGs monitored for Continuity and Connectivity or between an MEG monitored for Continuity and Connectivity and one MEG monitored only for Continuity.

CC-V OAM packets MUST be transmitted at a regular, operator's configurable, rate. The default CC-V transmission periods are application dependent (see section 5.1.4).

Proactive CC-V OAM packets are transmitted with the "minimum loss probability PHB" within a single network operator. This PHB is configurable on network operator's basis. PHBs can be translated at the network borders.

[Editor's note - Describe the relation between the previous paragraph and the fate sharing requirement. Need to clarify also in the requirement document that for proactive CC-V the fate sharing is related to the forwarding behavior and not to the QoS behavior]

In a bidirectional point-to-point transport path, when a MEP is enabled to generate pro-active CC-V OAM packets with a configured transmission rate, it also expects to receive pro-active CC-V OAM packets from its peer MEP at the same transmission rate. In a unidirectional transport path (either point-to-point or point-to-multipoint), only the source MEP is enabled to generate CC-V OAM packets and only the 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 pro-active CC-V information and forward these pro-active CC-V OAM packets as regular data packets.

To initialize the proactive CC-V monitoring on a configured ME without affecting traffic, the MEP source function (generating pro-active CC-V packets) should be enabled prior to the corresponding MEP sink function (detecting continuity and connectivity defects). When disabling the CC-V proactive functionality, the MEP sink function should be disabled prior to the corresponding MEP source function.

5.1.1. Defects identified by CC-V

Pro-active CC-V functions allow a sink MEP to detect the defect conditions described in the following sub-sections. For all of the described defect cases, the sink MEP SHOULD notify the equipment fault management process of the detected defect.

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 pro-active CC-V OAM packets from the peer MEP.

- o Entry criteria: if no pro-active CC-V OAM packets from the peer MEP (i.e. with the correct ME and peer MEP identifiers) are received within the interval equal to 3.5 times the receiving MEP's configured CC-V transmission period.

- o Exit criteria: a pro-active CC-V OAM packet from the peer MEP (i.e. with the correct ME and peer MEP identifiers) is received.

5.1.1.2. Mis-connectivity defect

When a pro-active CC-V OAM packet is received, a sink MEP identifies a mis-connectivity defect (e.g. mismerge or misconnection) with its

peer source MEP when the received packet carries an incorrect ME identifier.

o Entry criteria: the sink MEP receives a pro-active CC-V OAM packet with an incorrect ME ID.

o Exit criteria: the sink MEP does not receive any pro-active CC-V OAM packet with an incorrect ME ID for an interval equal at least to 3.5 times the longest transmission period of the pro-active CC-V OAM packets received with an incorrect ME ID since this defect has been raised. This requires the OAM message to self identify the CC-V periodicity as not all MEPs can be expected to have knowledge of all MEs.

5.1.1.3. MEP misconfiguration defect

When a pro-active CC-V packet is received, a sink MEP identifies a MEP misconfiguration defect with its peer source MEP when the received packet carries a correct ME Identifier but an unexpected peer MEP Identifier which includes the MEP's own MEP Identifier.

o Entry criteria: the sink MEP receives a CC-V pro-active packet with correct ME ID but with unexpected MEP ID.

o Exit criteria: the sink MEP does not receive any pro-active CC-V OAM packet with a correct ME ID and unexpected MEP ID for an interval equal at least to 3.5 times the longest transmission period of the pro-active CC-V OAM packets received with a correct ME ID and unexpected MEP ID since this defect has been raised.

5.1.1.4. Period Misconfiguration defect

If pro-active CC-V OAM packets are received with correct ME and MEP identifiers but with a transmission period different than its own configured transmission period, then a CC-V period mis-configuration defect is detected

o Entry criteria: a MEP receives a CC-V pro-active packet with correct ME ID and MEP ID but with a Period field value different than its own CC-V configured transmission period.

o Exit criteria: the sink MEP does not receive any pro-active CC-V OAM packet with a correct ME and MEP IDs and an incorrect transmission period for an interval equal at least to 3.5 times the longest transmission period of the pro-active CC-V OAM packets received with a correct ME and MEP IDs and an incorrect transmission period since this defect has been raised.

5.1.2. Consequent action

[editors: IMO this would be better folded into the specific defect types, If agreed I will edit accordingly]

A sink MEP that detects one of the defect conditions defined in section 5.1.1 MUST perform the following consequent actions. Some of these consequent actions SHOULD be enabled/disabled by the operator depending upon the application used (see section 5.1.4).

If a MEP detects an unexpected ME Identifier, or an unexpected MEP, it MUST block all the traffic (including also the user data packets) that it receives from the misconnected transport path.

If a MEP detects LOC defect and the CC-V monitoring is enabled 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 [12] recommends that CC-V proactive monitoring is enabled on every ME in order to reliably detect connectivity defects. However, CC-V proactive monitoring MAY be disabled by an operator on an ME. In the event of a misconnection between a transport path that is pro-actively monitored for CC-V and a transport path which 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 consequent wrong traffic delivering. For these reasons, the traffic block consequent action is applied even when a LOC condition occurs. This block consequent action MAY 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, an unexpected ME Identifier, or an unexpected MEP it MUST declare a signal fail condition at the transport path level.

If a MEP detects an Unexpected Period defect it SHOULD declare a signal fail condition at the transport path level.

[Editor's note - Transport equipment also performs defect correlation (as defined in G.806) in order to properly report failures to the transport NMS]. The current working assumption, to be further

investigated, is that defect correlations are outside the scope of this document and to be defined in ITU-T documents.]

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:

- o MEG ID; the MEG identifier to which the MEP belongs;

- o MEP-ID; the MEP's own identity inside the MEG;

- o list of peer MEPs inside the MEG. For a point-to-point MEG the list would consist of the single peer MEP ID from which the OAM packets are expected. In case of the root MEP of a p2mp MEG, the list is composed by all the leaf MEP IDs inside the ME. In case of the leaf MEP of a p2mp MEG, the list is composed by the root MEP ID (i.e. each leaf MUST know the root MEP ID from which it expect to receive the CC-V OAM packets).

- o transmission rate; the default CC-V transmission periods are application dependent (see section 5.1.4)

- o PHB; it identifies the per-hop behaviour of 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.

For statically provisioned transport paths the above information are statically configured; for dynamically established transport paths the configuration information are signaled via the control plane.

5.1.4. Applications for proactive CC-V

CC-V is applicable for fault management, performance monitoring, or protection switching applications.

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

- o Performance Monitoring: Performance monitoring 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.

o Protection Switching: in order to achieve sub-50ms the defect entry criteria should resolve in less than 50msec, and should budget sufficient portion of the 50 msec. to be available for consequent action processing. In some cases, when a slower recovery time is acceptable, it is also possible to lengthen the transmission rate.

It SHOULD be possible for the operator to configure these transmission rates for all applications, to satisfy his internal requirements.

In addition, the operator should be able to define the consequent action to be performed for each of these applications.

5.2. Remote Defect Indication

The Remote Defect Indication (RDI) is an indicator that is transmitted by a MEP to communicate to its peer MEPs that a signal fail condition exists. RDI is only used for bidirectional connections and is associated with proactive CC-V activation. The RDI indicator is piggy-backed onto the CC-V packet.

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. The RDI information will be included in all pro-active CC-V packets that it generates for the duration of the signal fail condition's existence.

[Editor's note - Add some forward compatibility information to cover the case where future OAM mechanisms that contributes to the signal fail detection (and RDI generation) are defined.]

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

MIPs as well as intermediate nodes not supporting MPLS-TP OAM are transparent to the RDI indicator and forward these proactive CC-V 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 defect condition clears, the MEP should clear the RDI indicator from subsequent transmission of pro-active CC-V packets. A MEP should clear the RDI defect upon reception of a pro-active CC-V packet from the source MEP with the RDI indicator cleared.

5.2.1. Configuration considerations

In order to support RDI indication, this may be a unique OAM message or an OAM information element embedded in a CV message. In this case the RDI transmission rate and PHB of the OAM packets carrying RDI should be the same as that configured for CC-V.

5.2.2. Applications for Remote Defect Indication

RDI is applicable for the following applications:

o Single-ended fault management - A MEP that receives an RDI indication from its peer MEP, can report a far-end defect condition (i.e. the peer MEP has detected a signal fail condition in the traffic direction from the MEP that receives the RDI indication to the peer MEP that has sent the RDI information).

o Contribution to far-end performance monitoring - The indication of the far-end defect condition is used as a contribution to the bidirectional performance monitoring process.

5.3. Alarm Reporting

The Alarm Reporting function relies upon an Alarm Indication Signal (AIS) message used to suppress alarms following detection of defect conditions at the server (sub-)layer.

o A server MEP that detects a signal fail conditions in the server (sub-)layer, can generate packets with AIS information in a direction opposite to its peers MEPs to allow the suppression of secondary alarms at the MEP in the client (sub-)layer.

A server MEP is responsible for notifying the MPLS-TP layer network MEP upon fault detection in the server layer network to which the server MEP is associated.

[editor: the above is confused. The server layer passes signal fail or whatever notification to the adaptation function which has knowledge of the client layer transport paths, otherwise we are discussing a layer violation. These may be MEP co-located end points or MIPs. It is the OAM functionality co-located with the adaptation function that performs AIS insertion into the client layer MPLS-TP paths.... If agreed I will re-word accordingly]

Only Server MEPs can issue MPLS-TP packets with AIS information. Upon detection of a signal fail condition the Server MEP can immediately start transmitting periodic packets with AIS information. These

periodic packets, with AIS information, continue to be transmitted until the signal fail condition is cleared. [editor: SEE ABOVE]

Upon receiving a packet with AIS information an MPLS-TP MEP detects an AIS defect condition and suppresses loss of continuity alarms associated with all of its peer MEPs. [editor: There can only be one MEP for the ME AIS has been received in association with] A MEP resumes loss of continuity alarm generation upon detecting loss of continuity defect conditions in the absence of AIS condition.

For example, let's consider a fiber cut between LSR 1 and LSR 2 in the reference network of Figure 3. Assuming that all the MEs described in Figure 3 have pro-active CC-V enabled, a LOC defect is detected by the MEPs of Sec12 SME, PSN13 LME, PW1 PTCME and PW1Z PME, however in transport network only the alarm associate to the fiber cut needs to be reported to NMS while all these 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 LSR2), LSR2 can generate the proper alarm in the physical layer and suppress the secondary alarm associated with the LOC defect detected on Sec12 SME. As both MEPs reside within the same node, this process does not involve any external protocol exchange. Otherwise, if the physical layer has not enough OAM capabilities to detect the fiber cut, the MEP of Sec12 SME in LSR2 will report a LOC alarm.

In both cases, the MEP of Sec12 SME in LSR 2 generates AIS packets on the PSN13 LME in order to allow its MEP in LSR3 to suppress the LOC alarm. LSR3 can also suppress the secondary alarm on PW1 PTCME because the MEP of PW1 PTCME resides within the same node as the MEP of PSN13 LME. The MEP of PW1 PTCME in LSR3 also generates AIS packets on PW1Z PME in order to allow its MEP in LSRZ to suppress the LOC alarm.

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

AIS packets are transmitted with the "minimum loss probability PHB" within a single network operator. This PHB is configurable on network operator's basis.

A MIP is transparent to packets with AIS information and therefore does not require any information to support AIS functionality.

5.4. Lock Reporting

To be incorporated in a future revision of this document

5.5. Packet Loss Monitoring

Packet Loss Monitoring (LM) is one of the capabilities supported by the MPLS-TP Performance Monitoring (PM) function in order to facilitate reporting of QoS information for a transport path. 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 bidirectional transport path) during the life time of the transport path. Each MEP performs measurements of its transmitted and received packets. These measurements are then transactionally correlated 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 queuing discipline 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).

5.5.1. Configuration considerations

In order to support proactive LM, the transmission rate and PHB associated with the LM OAM packets originating from a MEP need be configured as part of the LM provisioning procedures. LM OAM packets should be transmitted with the PHB that yields the lowest packet loss performance among the PHB Scheduling Classes or Ordered Aggregates (see RFC 3260 [15]) in the monitored transport path for the relevant network domain(s).

5.5.2. Applications for Packet Loss Monitoring

LM is relevant for the following applications:

o Single or double-end performance monitoring: determination of the packet loss performance of a transport path for Service Level Agreement (SLA) verification purposes.

o Single or double-end performance monitoring: determination of the packet loss performance of a PHB Scheduling Class or Ordered Aggregate within a transport path.

o Contribution to service unable time. Both near-end and far-end packet loss measurements contribute to performance metrics such as near-end severely errored seconds (Near-End SES) and far-end severely errored seconds (Far-End SES) respectively, which together contribute to unavailable time, in a manner similar to Recommendation G.826 [19] and Recommendation G.7710 [20].

5.6. Client Signal Failure Indication

The Client Signal Failure Indication (CSF) function 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) 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 transmission path does not support a native defect/alarm indication mechanism, e.g. FDI/AIS.

A source MEP starts transmitting a CSF indication to its peer MEP when it receives a local client signal defect notification via its local CSF function. Mechanisms to detect local client signal fail defects are technology specific.

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

Either there needs to be a 1:1 correspondence between the client and the ME, or when multiple clients are multiplexed over a transport path, the CSF message requires additional information to permit the client instance to be identified.

5.6.1. Configuration considerations

In order to support CSF indication, the CSF transmission rate and PHB of the CSF OAM message/information element should be configured as part of the CSF configuration.

5.6.2. Applications for Client Signal Failure Indication

CSF is applicable for the following applications:

o Single-ended fault management - A MEP that receives a CSF indication from its peer MEP, can report a far-end client defect condition (i.e. the peer MEP has been informed of local client signal fail condition in the traffic direction from the client to the peer MEP that transmitted the CSF).

o Contribution to far-end performance monitoring - The indication of the far-end defect condition may be used to account on network operator contribution to the bidirectional performance monitoring process.

CSF supports the application described in Appendix VIII of ITU-T G.806 [18].

5.7. Delay Measurement

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. Specifically, pro-active 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 bidirectional transport path) during a configurable time interval.

Pro-active DM can be operated in two ways:

o One-way: a MEP sends 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 synchronized precision time at either MEP by means outside the scope of this framework.

o Two-way: a MEP sends 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 containing all the required information to facilitate two-way packet delay and/or two-way packet delay variation measurements from the viewpoint of the source MEP.

5.7.1. Configuration considerations

In order to support pro-active DM, the transmission rate and PHB associated with the DM OAM packets originating from a MEP need be configured as part of the DM provisioning procedures. DM OAM packets

should be transmitted with the PHB that yields the lowest packet loss performance among the PHB Scheduling Classes or Ordered Aggregates (see RFC 3260 [15]) in the monitored transport path for the relevant network domain(s).

5.7.2. Applications for Delay Measurement

DM is relevant for the following applications:

- Single or double-end performance monitoring: determination of the delay performance of a transport path for SLA verification purposes.

- Single or double-end performance monitoring: determination of the delay performance of a PHB Scheduling Class or Ordered Aggregate within a transport path

6. OAM Functions for on-demand monitoring

[Munich: Note the fwk needs to be explicit about the mapping of functions to the tools we have chosen.]

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

[editor: we would have to babysit a lot fewer words if we folded this into section 5 and simply indicated which transactions existed in both proactive and reactive forms... if agreed I will edit accordingly]

6.1. Connectivity Verification

In order to preserve network resources, e.g. bandwidth, processing time at switches, it may be preferable to not use proactive CC-V. In order to perform fault management functions, network management may invoke periodic on-demand bursts of on-demand CV packets.

Use of on-demand CV is dependent on the existence of a bi-directional connection ME, because it requires the presence of a return path in the data plane.

[Editor's note - Clarify in the sentence above and within the paragraph that on-demand CV requires a return path to send back the reply to on-demand CV packets]

An additional use of on-demand CV would be to detect and locate a problem of connectivity when a problem is suspected or known 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 ME that is being checked. The on-demand functionality may be used to check either an entire ME (end-to-end) or between a MEP to a specific MIP. This functionality may not be available for associated bidirectional paths as the MIP may not have a return path to the source MEP for the on-demand CV transaction.

On-demand CV may generate a one-time burst of on-demand CV packets, or be used to invoke periodic, non-continuous, bursts of on-demand CV packets. The number of packets generated in each burst is configurable at the MEPs, and should take into account normal packet-loss conditions.

When invoking a periodic check of the ME, the source MEP should issue a burst of on-demand CV packets that uniquely identifies the ME being verified. The number of packets and their transmission rate should be pre-configured and known to both the source MEP and the target MEP or MIP. The source MEP should use the TTL field to indicate the number of hops necessary, when targeting a MIP and use the default value when performing an end-to-end check [IB => This is quite generic for addressing packets to MIPs and MEPs so it is better to move this text in section 2]. 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 source MEP, the LOC defect state is entered.

[Editor's note - We need to add some text for the usage of on-demand CV with different packet sizes, e.g. to discover MTU problems.]

When a connectivity problem is detected (e.g. via a proactive CC-V OAM tool), an on-demand CV tool can be used to check the path. The series should check CV from MEP to peer MEP on the path, and if a fault is discovered, by lack of response, then additional checks may be performed to each of the intermediate MIP to locate the fault.

[Dave: this seems a bit warped as the original discussion was about not spending resources on proactive CC-V, so can we just be honest about "when the incredibly pissed off customer calls, an on demand CV tool..."]

6.1.1. Configuration considerations

For on-demand CV the MEP should support the configuration of the number of packets to be transmitted/received in each burst of transmissions and their packet size. The transmission rate should be configured between the different nodes.

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.

The PHB of the on-demand CV packets should be configured as well.

[Editor's note - We need to be better define the reason for such configuration]

6.2. Packet Loss Monitoring

On-demand Packet Loss (LM) is one of the capabilities supported by the MPLS-TP Performance Monitoring function in order to facilitate diagnostic of QoS performance for a transport path. As proactive LM, on-demand 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.

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 bidirectional transport path) during a pre-defined monitoring period. Each MEP performs measurements of its transmitted and received packets. These measurements are then correlated evaluate the packet loss performance metrics of the transport path. [Dave: again are we discussing simply discard eligibility and no other PHB impacts?]

6.2.1. Configuration considerations

In order to support on-demand LM, the beginning and duration of the LM procedures, the transmission rate and PHB associated with the LM OAM packets originating from a MEP must be configured as part of the on-demand LM provisioning procedures. LM OAM packets should be transmitted with the PHB that yields the lowest packet loss performance among the PHB Scheduling Classes or Ordered Aggregates (see RFC 3260 [15]) in the monitored transport path for the relevant network domain(s).

6.2.2. Applications for On-demand Packet Loss Monitoring

On-demand LM is relevant for the following applications:

- o Single-end performance monitoring: diagnostic of the packet loss performance of a transport path for SLA trouble shooting purposes.

- o Single-end performance monitoring: diagnostic of the packet loss performance of a PHB Scheduling Class or Ordering Aggregate within a transport path for QoS trouble shooting purposes.

6.3. Diagnostic

To be incorporated in a future revision of this document

[Munich: Need to describe the two types of loopback - LBM/LBR and traffic loopback enhanced with variable sized packets in the on demand cases.

One objective of diags is fault location, we need to make clear how we apply the tools to fault location.

At the top of each section we need to describe the detailed requirements and then in the rest of the section describe how it is met.]

6.4. Route Tracing

After e.g. provisioning an MPLS-TP LSP or for trouble shooting purposes, it is often necessary to trace a route covered by an ME from a source MEP to the sink MEP including all the MIPs in-between. The route tracing function is providing this functionality. Based on the fate sharing requirement of OAM flows, i.e. OAM packets receive the same forwarding treatment as data packet, route tracing is a basic means to perform CV and, to a much lesser degree, CC. 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 trace could be implemented e.g. by an MPLS traceroute-like function [RFC4379]. However, route tracing should always return the full list of MIPs and the peer MEP in it answer(s). In case a defect exist, the route trace function needs to be able to detect it and stop automatically returning the incomplete list of OAM entities that it was able to trace.

The configuration of the route trace function must at least support the setting of the trace depth (number of hops)_and the number of trace attempts before it gives up. Default setting need to be configurable by the operator, too.

6.5. Delay Measurement

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. 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 bidirectional transport path) during a configurable time interval.

On-demand DM can be operated in two ways:

- o One-way: a MEP sends 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.

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

6.5.1. Configuration considerations

In order to support on-demand DM, the beginning and duration of the DM procedures, the transmission rate and PHB associated with the DM OAM packets originating from a MEP need be configured as part of the LM provisioning procedures. DM OAM packets should be transmitted with the PHB that yields the lowest packet delay performance among the PHB Scheduling Classes or Ordering Aggregates (see RFC 3260 [15]) in the monitored transport path for the relevant network domain(s).

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 of the on-demand OAM DM packet.

6.5.2. Applications for Delay Measurement

DM is relevant for the following applications:

- o Single or double-end performance monitoring: determination of the packet delay and/or delay variation performance of a transport path for SLA verification purposes.

- o Single or double-end performance monitoring: determination of the packet delay and/or delay variation a PHB Scheduling Class or Ordering Aggregate within a transport path

- o Contribution to service unable time. Packet delay measurements may contribute to performance metrics such as near-end severely errored seconds (Near-End SES) and far-end severely errored seconds (Far-End SES), which together contribute to unavailable time.

6.6. Lock Instruct

To be incorporated in a future revision of this document

7. Security Considerations

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

OAM traffic can reveal sensitive information such as passwords, performance data and details about e.g. the network topology. The nature of OAM data therefore suggests to have some form of authentication, authorization and encryption 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.

Mechanisms that the framework does not specify might be subject to additional security considerations.

8. IANA Considerations

No new IANA considerations.

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

T-MPLS Ad Hoc Group in ITU-T) involved in the definition and specification of MPLS Transport Profile.

This document was prepared using 2-Word-v2.0.template.dot.

10. References

10.1. Normative References

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

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

[3] Rosen, E., et al., "MPLS Label Stack Encoding", RFC 3032, January 2001

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

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

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

[7] Bocci, M., Bryant, S., "An Architecture for Multi-Segment Pseudo Wire Emulation Edge-to-Edge", draft-ietf-pwe3-ms-pw-arch-05 (work in progress), September 2008

[8] Bocci, M., et al., "A Framework for MPLS in Transport Networks", draft-ietf-mpls-tp-framework-01 (work in progress), June 2009

[9] Vigoureux, M., Bocci, M., Swallow, G., Ward, D., Aggarwal, R., "MPLS Generic Associated Channel", RFC 5586, June 2009

[10] Swallow, G., Bocci, M., "MPLS-TP Identifiers", draft-swallow-mpls-tp-identifiers-01 (work in progress), July 2009

10.2. Informative References

[11] Niven-Jenkins, B., Brungard, D., Betts, M., sprecher, N., Ueno, S., "MPLS-TP Requirements", RFC 5654, September 2009

[12] Vigoureux, M., Betts, M., Ward, D., "Requirements for OAM in MPLS Transport Networks", draft-ietf-mpls-tp-oam-requirements-03 (work in progress), August 2009

[13] Sprecher, N., Nadeau, T., van Helvoort, H., Weingarten, Y., "MPLS-TP OAM Analysis", draft-sprecher-mpls-tp-oam-analysis-04 (work in progress), May 2009

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

[15] Grossman, D., "New terminology and clarifications for Diffserv", RFC 3260, April 2002.

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

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

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

[19] ITU-T Recommendation G.826 (12/02), "End-to-end error performance parameters and objectives for international, constant bit-rate digital paths and connections", December 2002

[20] ITU-T Recommendation G.7710 (07/07), "Common equipment management function requirements", July 2007

[21] ITU-T Recommendation Y.2611 (06/12), " High-level architecture of future packet-based networks", 2006

Authors' Addresses

Dave Allan (Editor)
Ericsson

Email: david.i.allan@ericsson.com

Italo Busi (Editor)
Alcatel-Lucent

Email: Italo.Busi@alcatel-lucent.it

Ben Niven-Jenkins (Editor)
BT

Email: benjamin.niven-jenkins@bt.com

Contributing Authors' Addresses

Annamaria Fulignoli
Ericsson

Email: annamaria.fulignoli@ericsson.com

Enrique Hernandez-Valencia
Alcatel-Lucent

Email: enrique@alcatel-lucent.com

Lieven Levrau
Alcatel-Lucent

Email: llevrau@alcatel-lucent.com

Dinesh Mohan
Nortel

Email: mohand@nortel.com

Vincenzo Sestito
Alcatel-Lucent

Email: vincenzo.sestito@alcatel-lucent.it

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.fr

Yaacov Weingarten
Nokia Siemens Networks

Email: yaacov.weingarten@nsn.com

Rolf Winter
NEC

Email: Rolf.Winter@nw.neclab.eu