Network Working Group
Internet-Draft
Intended status: Informational
Expires: October 31, 2013

Y. Weingarten

S. Bryant
Cisco
D. Ceccarelli
D. Caviglia
F. Fondelli
Ericsson
M. Corsi
Altran
B. Wu
X. Dai
ZTE Corporation

April 29, 2013

Applicability of MPLS-TP Linear Protection for Ring Topologies
draft-ietf-mpls-tp-ring-protection-06.txt

Abstract

This document presents an applicability of existing MPLS protection mechanisms, both local and end-to-end, to Multi-Protocol Label Switching Transport Profile (MPLS-TP) in ring topologies. This document does not propose any new mechanisms or protocols. Protection on rings offers a number of opportunities for optimization as the protection choices are starkly limited (all traffic traveling one way around a ring can only be switched to travel the other way on the ring), but also suffers from some complications caused by the limitations of the topology.

Requirements for MPLS-TP protection especially for protection in ring topologies are discussed in "Requirements of an MPLS Transport Profile" (RFC 5654) and "MPLS Transport Profile (MPLS-TP) Survivability Framework" (RFC 6372). This document shows how MPLS-TP linear protection as defined in RFC 6378 can be applied to single ring topologies, discusses how most of the requirements are met, and describes scenarios in which the function provided by applying linear protection in a ring topology falls short of some of the requirements.

This document is a product of a joint Internet Engineering Task Force (IETF) / International Telecommunications Union Telecommunications 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.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

Multi-Protocol Label Switching Transport Profile (MPLS-TP) has been standardized as part of a joint effort between the Internet Engineering Task Force (IETF) and the International Telecommunication Union Standardization (ITU-T). These specifications are based on the requirements that were generated from this joint effort.

The MPLS-TP requirement document [RFC5654] includes a requirement to support a network that may include sub-networks that constitute an MPLS-TP ring as defined in the document. However, the document does not identify any protection requirements specific to a ring topology. However, the requirements state that specific protection mechanisms applying to ring topologies may be developed if these allow the network to minimize:

- Number of OAM entities needed to trigger the protection
- Number of elements of recovery needed
- Number of labels required
- Number of control and management plane transactions during a maintenance operation
- Impact of signaling and routing information exchanged during protection, in the presence of a control plane

This document describes how applying a set of basic MPLS-TP linear protection mechanisms defined in [RFC6378] can be used to provide protection of the data flows that traverse an MPLS-TP ring. These mechanisms provide data flow protection due to any switching trigger within a reasonable time frame and optimize the criteria set out in [RFC5654], as summarized above. This document does not define any new protocol mechanisms or procedures.

A related topic in [RFC5654] addresses the required support for interconnected rings. This topic involves various scenarios that require further study and will be addressed in a separate document, based on the principles outlined in this document.

1.1. Problem statement

Ring topologies, as defined in [RFC5654], are used in transport networks. When designing a protection mechanism for a single ring topology, there is a need to address both -

1. A point-to-point transport path that either originates at or enters an MPLS-TP capable ring at one node, the ingress node, and exits the ring at a single egress node possibly continuing beyond the ring.

2. Where the ring is being used as a branching point for a point-to-multipoint transport path, i.e. the transport path originates at or enters the MPLS-TP capable ring at the ingress node and exits through a number of egress nodes, possibly continuing beyond the ring.

In either of these two situations, there is a need to address the following different cases -

1. One of the ring links causes a fault condition. This could be either a unidirectional or bidirectional fault, and should be detected by the neighboring nodes.

2. One of the ring nodes causes a fault condition. This condition is invariably a bidirectional fault (although in rare cases of misconfiguration this could be detected as a unidirectional fault) and should be detected by the two neighboring ring nodes.

3. An operator command changes the operational state of a node or a link, or specifically triggers a protection action is issued to a specific ring node. A description of the different operator commands is found in Section 4.13 of [RFC4427]. Examples of these commands include Manual Switch, Forced Switch, or Clear operations.

The protection domain addressed in this document is limited to the traffic that traverses on the ring. Protection triggers on the transport path prior to the ring ingress node or beyond the egress nodes may be protected by some other mechanism.

1.2. Scope of the document

This document addresses the requirements that appear in Section 2.5.6.1 of [RFC5654] on Ring Protection based on the application of the linear protection as defined in [RFC6378]. Requirement R93 regarding the support of interconnected rings and protection of faults in the interconnection nodes and links is for further study.

In addition, requirement R105 requiring the support of lockout of specific nodes or spans is only supported to the degree that it is supported by the linear protection mechanism.

1.3. Terminology and Notation

The terminology used in this document is based on the terminology defined in the MPLS-TP framework documents:

- MPLS-TP Framework[RFC5921]
- MPLS-TP OAM Framework[RFC6371]
- MPLS-TP Survivability Framework[RFC6372]

The MPLS-TP Framework document [RFC5921] defines a Sub-Path Maintenance Entity (SPME) construct that can be defined between any two Label Switching Routers (LSR) of an MPLS-TP Label Switched Path (LSP). This SPME may be configured as a co-routed bidirectional path. The SPME is defined to allow management and monitoring of any segment of a transport path. This concept will be used extensively throughout the document to support protection of the traffic that traverses an MPLS-TP ring.

In addition, we describe the use of the label stack in connection with the redirecting of data packets by the protection mechanism. The following syntax will be used to describe the contents of the label stack:

1. The label stack will be enclosed in square brackets ("[]")
2. Each level in the stack will be separated by the '|' character. It should be noted that the label stack may contain additional levels however, we only present the levels that are germane to the protection mechanism.
3. When applicable, the S-bit (signifying that a given label is the bottom of the label stack) will be denoted by the string '+S' within the label. If a label is not shown with '+S' that label may or may not be the bottom label in the stack. '+S' is only shown when it is important to illustrate that a given label is definitely the last one in the label stack.
4. The label of the LSP at the ingress point to the ring will be denoted by the string "LI" and the label of the LSP that is expected at the egress point from the ring will be denoted by the string "LE", and "LSE" will denote the label expected at the exit LSR of a SPME (if it is different from the egress point from the ring, for example as described in Section 2.3).
5. The label for a SPME will be denoted by Pxi(y) where x and y are LSR identifiers and the intention is to the label for LSR-x to

transmit to LSR-y over the SPME whose index is i.

For example -

- the label stack [LI] denotes the label stack received at the ingress node of the ring. This may have additional labels after LI, e.g. a PW label however, this is irrelevant to the discussion of the protection scenario.

- [PB1(G)|LE] denotes a stack whose top-label is the SPME-1 label for LSR-B to transmit the data packet to LSR-G, the second label is the label that would be used by the egress LSR to continue the packet on the original LSP.

- If "LE" were the bottom label in the stack, then the label stack would be shown as [PB1(G)|LE+S].

1.4. Contributing Authors

The authors would like to acknowledge the following individuals that contributed their insights and advice to this work:

Nurit Sprecher (NSN)

Akira Sakurai (NEC)

Rolf Winter (NEC)

Eric Osborne (Cisco)

2. Point-to-point (P2P) Ring Protection

There are two protection architecture mechanisms, that have historically been applied to ring topologies, based on SDH specifications [G.841], and have been proposed in various forums to perform recovery of a topological ring network - "Wrapping" and "Steering". The following sub-sections examine these two mechanisms, as applied to an MPLS transport network.

2.1. Wrapping

Wrapping is defined as a local protection architecture. This mechanism is local to the nodes that are neighbors to the detected fault. When a fault is detected (either a link or node failure), the neighboring node can identify that the fault would prevent forwarding of the data along the data path. Therefore, in order to continue the data along the path, there is need to "wrap" all data traffic around

the ring, on an alternate data path, until arriving at the node that is on the opposite side of the fault. When this far-side node also detects that there is a fault condition on the working path, it can identify that the data traffic that is arriving on the alternate (protecting) data path is intended for the "broken" data path. Therefore, again taking a local decision, can wrap the data back onto the normal working path until the egress from the ring segment.

Wrapping behavior is similar to MPLS-TE FRR as defined in [RFC4090] using either bypass or detour tunnels. Applying this methodology to MPLS, it is possible to wrap the traffic of each LSP around the failed links via a detour tunnel using a different label for each LSP or to wrap all LSPs using a bypass tunnel and a single label.

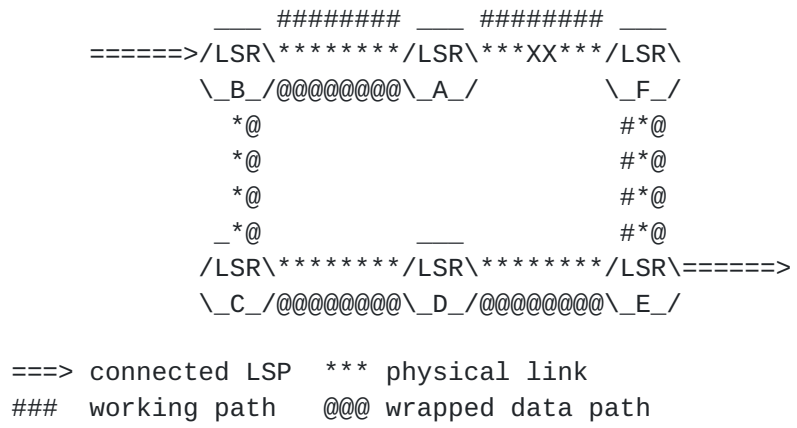

Figure 1: Wrapping protection for P2P path

Consider the LSP that is shown in Figure 1 that enters the ring of LSRs at LSR-B and exits at LSR-E. The normal working path LSP follows through LSRs B-A-F-E. If a fault is detected on the link A<->F, then the wrapping mechanism decides that LSR-A would wrap the traffic around the ring, on a wrapped data path A-B-C-D-E-F, to arrive at LSR-F (on the far side of the failed link). LSR-F would then wrap the data packets back onto the working path F->E to the egress node. In this protection scheme, the traffic will follow the path - B-A-B-C-D-E-F-E.

This protection scheme is simple in the sense that there is no need for coordination between the different LSR in the ring - only the LSRs that detect the fault must wrap the traffic, either onto the wrapped data path (at the near-end) or back to the working path (at the far-end). However, coordination of the switchover to the protection path would be needed to maintain the traffic on a co-routed bidirectional LSP even in cases of a unidirectional fault condition.

The following considerations should be taken into account when considering use of wrapping protection:

- Detection of loss-of-continuity or mis-connectivity should be performed at the link level and/or per LSR when using node-level protection. Configuration of the protection being performed (i.e. link protection or node protection) needs to be performed a-priori, since the configuration of the proper protection path is dependent upon this decision.

- There is a need to define a data-path that traverses the alternate path around the ring to connect between the two neighbors of the detected fault. If protecting both the links and the nodes of a LSP, then, for a ring with N nodes, there is a need for O(2N) alternate paths.

- When wrapping, the data is transmitted over some of the links twice, once in each direction. For example, in the figure above the traffic is transmitted both B->A and then A->B, later it is transmitted E->F and F->E. This means that there is additional bandwidth needed for this protection.

- If a double-fault situation occurs in the ring, then wrapping will not be able to deliver any packets except between the ingress and the first fault location encountered on the working path. This is based on the need for wrapping to connect between the neighbors of the fault location, and this is not possible in the segmented ring.

- The resource pre-allocation for all of the alternate-paths could be problematic [causing massive over subscription of the available resources]. However, since most of these alternate paths will not be used simultaneously, there is the possibility of allocating '0' resources and depend on the NMS to allocate the proper resources around the ring, based on actual traffic usage.

- Wrapping also involves a small increase in traffic latency in delivering the packets, as a result of traversing the entire ring, during protection.

2.2. Steering

The second common scheme for ring protection, steering, takes advantage of the ring topology by defining two paths from the ingress point (to the ring) to the egress point going in opposite directions around the ring. This is illustrated in Figure 2, where if we assume that the traffic needs to enter the ring from node B and exit through node F, we could define a primary path through nodes B-A-F, and an

alternate path through the nodes B-C-D-E-F. In steering the switching is always performed by the ingress node (node B in Figure 2). If a fault condition is detected anywhere on the working path (B-A-F), then the traffic would be redirected by B to the alternate path (i.e. B-C-D-E-F).

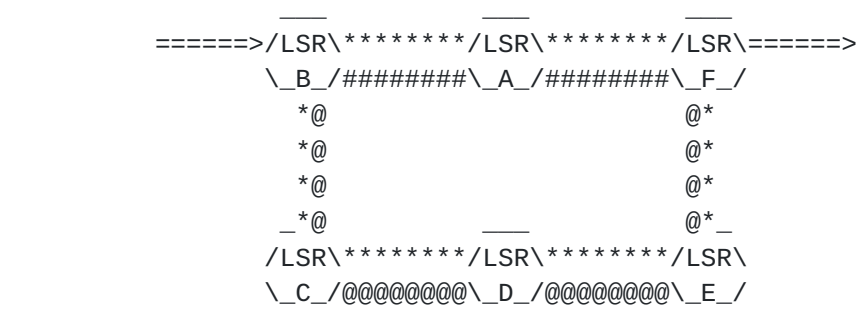

Figure 2: Steering protection in an MPLS-TP ring

This mechanism bears similarities to linear 1:1 protection [RFC6372]. The two paths around the ring act as the working and protection paths. There is need to communicate to the ingress node the need to switch over to the protection path and there is a need to coordinate the switchover between the two end-points of the protected domain.

The following considerations must be taken into account regarding the steering architecture:

- Steering relies on a failure detection method that is able to notify the ingress node of the fault condition. This may involve different OAM functionality described in [RFC6371], e.g. Remote Defect Indication, Alarm reporting.

- The process of notifying the ingress node adds to the latency of the protection switching process, after the detection of the fault condition.

- While there is no need for double bandwidth for the data path, there is the necessity for the ring to maintain enough capacity for all of the data in both directions around the ring.

2.3. SPME for P2P protection of a ring topology

The SPME concept was introduced by [RFC5921] to support management and monitoring an arbitrary segment of a transport. However, an SPME is essentially a valid LSP that may be used to aggregate all LSP

traffic that traverses the sub-path delineated by the SPME. An SPME may be monitored using the OAM mechanisms as described in the MPLS-TP OAM Framework document [RFC6371].

When defining an MPLS-TP ring as a protection domain, there is a need to design a protection mechanism that protects all the LSPs that cross the MPLS-TP ring. For this purpose, we associate a (working) SPME with the segment of the transport path that traverses the ring. In addition, we configure an alternate (protecting) SPME that traverses the ring in the opposite direction around the ring. The exact selection of the SPMEs is dependent on the type of transport path and protection that is being implemented and will be detailed in the following sub-sections.

Based on this architectural configuration for protection of ring topologies, it is possible to limit the number of alternate paths needed to protect the data traversing the ring. In addition, since we will perform all of the OAM functionality on the SPME configured for the traffic, we can minimize the number of OAM sessions needed to monitor the data traffic of the ring - rather than monitoring each individual LSP.

In all of the following subsections, we use 1:1 linear protection [RFC6372] [RFC6378] to perform protection switching and coordination when a signal fault is detected. The actual configuration of the SPMEs used may change dependent upon the choice of methodology and this will be detailed in the following sections. However, in all of these configurations the mechanism will be to transmit the data traffic on the primary SPME, while applying OAM functionality over both the primary and the secondary SPME to detect signal fault conditions on either path. If a signal fault is detected on the primary SPME, then the mechanism described in [RFC6378] shall be used to coordinate a switch-over of data traffic to the secondary SPME.

Assuming that the SPME is implemented as an hierarchical LSP, packets that arrive at LSR-B with a label stack [LI] will have the SPME label pushed at LSR-B and the LSP label will be swapped for the label that is expected by the egress LSR (i.e. the packet will arrive at LSR-A with a label stack of [PA1(B)|LE], arrive at LSR-F with [PE1(F)|LE]). The SPME label will be popped by LSR-F and the LSP label will be treated appropriately at LSR-F and forwarded along the LSP, outside the ring. This scenario is true for all LSP that are aggregated by this primary SPME.

2.3.1. Path SPME for Steering

A P2P SPME that traverses part of a ring has two Maintenance Entity Group End Points (MEPs), each one acts as the ingress and egress in

one direction of the bidirectional SPME. Since the SPME is traversing a ring we can take advantage of another characteristic of a ring - there is always an alternative path between the two MEPs, i.e. traversing the ring in the opposite direction. This alternative SPME can be defined as the protection path for the working path that is configured as part of the LSP and defined as a SPME.

For each pair of SPMEs that are defined in this way, it is possible to verify the connectivity and continuity by applying the MPLS-TP OAM functionality to both the working and protection SPME. If a discontinuity or mis-connectivity is detected then the MEPs will become aware of this condition, and could perform a protection switch of all LSPs to the alternate, protection SPME.

The following figure shows an MPLS-TP ring that is part of a larger MPLS-TP network. The ring could be used as a network segment that may be traversed by numerous LSPs. In particular, the figure shows that for all LSPs that connect to the ring at LSR-B and exit the ring from LSR-F, we configure two SPME through the ring (the first SPME traverses along B-A-F, and the second SPME traverses B-C-D-E-F).

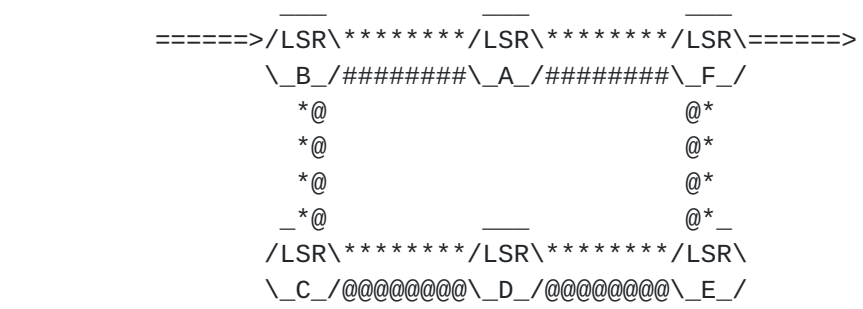

Figure 3: An MPLS-TP ring

This protection mechanism is identical to application of 1:1 linear protection[RFC6372] [RFC6378] to the pair of SPMEs. Under normal conditions, all LSP data traffic will be transmitted on the working SPME. If the linear protection is triggered, by either the OAM indication, an other fault indication trigger, or an operator command, then the MEPs will select the protection SPME to transmit all LSP data packets.

The protection SPME will continue to transmit the data packets until the stable recovery of the fault condition. Upon recovery, i.e. the fault condition has cleared and the network is stabilized, the ingress LSR could switch traffic back to the working SPME, if the

protection domain is configured for revertive behavior.

The control of the protection switching, especially for cases of operator commands, would be covered by the protocol defined in [RFC6378].

2.3.2. Wrapping link protection with segment based SPME

It is possible to use the SPME mechanism to perform segment-based protection. For each link in the ring, we define two SPME - the first is a SPME between the two LSRs that are connected by the link, and the second SPME between these same two LSRs but traversing the entire ring (except the link that connects the LSRs). In Figure 4 we show the primary SPME that connects LSR-A & LSR-F over a segment connection, and the secondary SPME that connects these same LSRs by traversing the ring in the opposite direction.

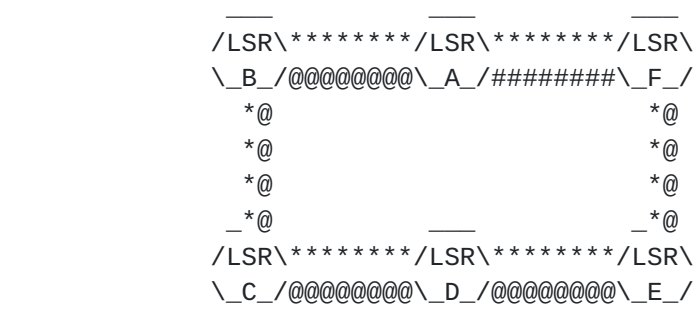

Figure 4: Segment SPMEs

By applying OAM monitoring of these two SPME (at each LSR), it is possible to affect a wrapping protection mechanism for the LSP traffic that traverses the ring. The LSR on either side of the segment would identify that there is a fault condition on the link and redirect all LSP traffic to the secondary SPME. The traffic would traverse the ring until arriving at the neighboring (relative to the segment) LSR. At this point, the LSP traffic would be redirected onto the original LSP, quite likely over the neighboring SPME.

Following the progression of the label stack through this switching operation (for a LSP that enters the ring at LSR B and exits the ring at LSR E):

1. The data packet arrives at LSR-A with label stack [L1+S] (i.e. top label from the LSP and bottom-of-stack indicator)

2. In the normal case (no protection switching), LSR-A forwards the packet with label stack [PA1(F)|LSE+S] (i.e. swap the label for the LSP, to be acceptable to the SPME egress, and push the label for the primary SPME from LSR-A to LSR-F).

3. When protection switching is in-effect, LSR-A forwards the packet with label stack [PA2(B)|LSE+S] (i.e. LSR-A pushed the label for the secondary SPME from LSR-A to LSR-F, after swapping the label of the lower level LSP). This will be transmitted along the secondary SPME until LSR-E forwards it to LSR-F with label stack [PE2(F)|LSE+S].

4. When the packet arrives at LSR-F, it will pop the SPME label, process the LSP label, and forward the packet to the next point, possibly pushing a SPME label if the next segment is likewise protected.

2.3.3. Wrapping node protection

Implementation of protection at the node level would be similar to the mechanism described in the previous sub-section. The difference would be in the SPMEs that are used. For node protection, the primary SPME would be configured between the two LSR that are connected to the node that is being protected (see SPME between LSR-A and LSR-E through LSR-F in Figure 5), and the secondary SPME would be configured between these same nodes, going around the ring (see secondary SPME in Figure 5).

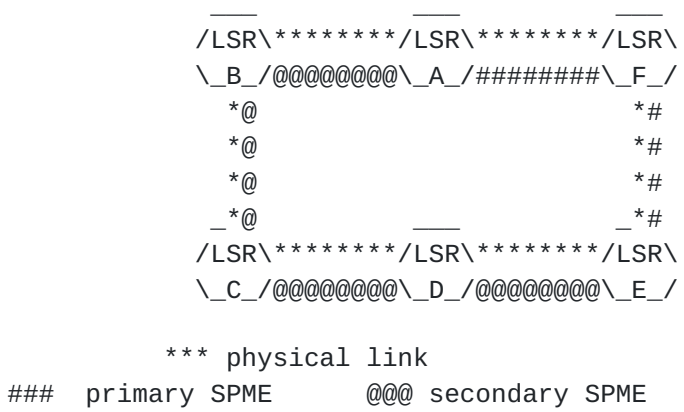

Figure 5: Node-protection SPMEs

The protection mechanism would work similarly - based on 1:1 linear

protection [RFC6372], triggered by OAM functions on both SPMEs, and wrapping the data packets onto the secondary SPME at the ingress MEP (e.g. LSR-A in the figure) of the SPME and back onto the continuation of the LSP at the egress MEP (e.g. LSR-E in the figure) of the SPME.

2.3.4. Wrapping for link and node protection

In the different types of wrapping presented in Section 2.3.2 and Section 2.3.3, there is a limitation that the protection mechanism must a priori decide whether it is protecting for link or node failure. In addition, the neighboring LSR, that detects the fault, cannot readily differentiate between a link failure or a node failure.

It would be possible to configure extra SPME to protect both for link and node failures, arriving at a configuration of the ring that is shown in Figure 6. Here there are three protection SPME configured:

- Secondary node#1 would be used to divert traffic as a result of an indication that LSR-F is not available, it redirects traffic to be transmitted between LSR-A and LSR-E.

- Secondary node#2 would be used to divert traffic as a result of an indication that LSR-A is not available, it redirects traffic to be transmitted between LSR-F and LSR-B.

- Secondary segment would be used to divert traffic as a result of an indication that the segment between LSR-A and LSR-F is not available, it redirects traffic to be transmitted between LSR-A and LSR-F on the long circuit of the ring.

Choosing the SPME to use for the wrapping would, however, then involve considerable effort and could result in the protected traffic not sharing the same protection path in both directions.

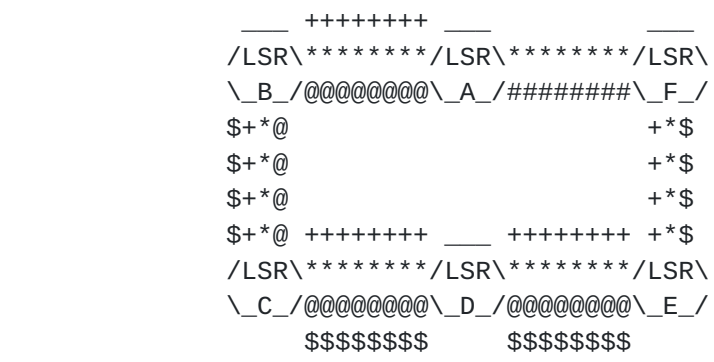

Figure 6: Segment & Node protection SPMEs

2.4. Analysis of P2P protection

Analyzing the mechanisms described in the above subsections we can point to the following observations (based on a ring with N nodes, assumed to be not more than 16):

- Number of SPME that need to be configured - for steering SPME protection (Section 2.3.1) = O(2N^2) [two SPME from each ingress LSR to each other node in the ring], for wrapping based on SPME either as described in Section 2.3.2 and Section 2.3.3 = O(2N) [however, the operator must decide a priori on whether to protect for link failures or node failures at each point]

- Number of OAM sessions at each node - for steering = O(2N), for SPME wrapping = 3

- Bandwidth requirements - for SPME-based steering: single bandwidth at each link, for wrapping: double bandwidth at links that are between ingress and wrapping node and between second wrapping node and egress.

- Special considerations - for SPME based steering: latency of OAM detection of fault condition by ingress MEP [using Alarm-reporting could optimize over using CC-V only], for SPME wrapping: at each node must decide a priori whether protecting for link or node failures. To protect for both node and link failures would increase the complexity of deciding which protection path to use, as well as, violating the co-routedness of the protected traffic.

Based on this analysis, using steering as described in Section 2.3.1 would be the recommended protection mechanism due to its simplicity.

It should be pointed out that the number of SPME involved in this protection could be reduced by eliminating SPME between pairs of LSR that are not used as an ingress and egress pair.

2.4.1. Recommendations for protection of P2P paths traversing a ring

Based on the analysis presented, while applying linear protection to effect Wrapping protection to a ring topology is possible as demonstrated, this does have certain limitations in addressing some of the required behavior. The limitations include:

- Need to a-priori configure the protection for link or node protection
- Increased number of SPME that need to be defined
- Difficulty in addressing cases of multiple failures in the ring

Application of linear protection, based on the use of SPME within the ring, to implement a Steering methodology to protect a ring topology is rather straight forward, overcomes the limitations listed above, and scales very well. For this and other reasons listed previously, the authors recommend the use of Steering to provide protection of a ring topology when using the mechanisms described in this document for protection of P2P paths that traverse the ring.

3. Point-to-multipoint protection

[RFC5654] requires that ring protection must provide protection for unidirectional point-to-multipoint paths through the ring. Ring topologies provide a ready platform for supporting such data paths. A Point-to-multipoint (P2MP) LSP in an MPLS-TP ring would be characterized by a single ingress LSR and multiple egress LSRs. The following sub-sections will present methods to address the protection of the ring-based sections of these LSP.

3.1. Wrapping for P2MP LSP

When protecting a P2MP ring data path using the wrapping architecture, the basic operation is similar to the description given, as the traffic has been wrapped back onto the normal working path on the far-side of the detected fault and will continue to be transported to all of the egress points.

It is possible to optimize the performance of the wrapping mechanism when applied to P2MP LSPs by exploiting the topology of ring networks.

This improved mechanism, which we call Ring Optimized Multipoint Wrapping (ROM-Wrapping), behaves much the same as classical wrapping. However, ROM-Wrapping configures protection P2MP LSP, relative to each node that is considered a failure risk, from the upstream node and all egress nodes (for the particular LSP) downstream from the failure risk.

Referring to Figure 7, it is possible to identify the protected (working) LSP (A-B-{C}-{D}-E-{F}) and one possible backup (protection) LSP (note:the egress nodes are indicated by the curly braces). This protection LSP will be used to wrap the data back around the ring to protect against a failure on link B-C. This protection LSP is also a P2MP LSP that is configured with egress points (at nodes F, D, & C) complementary to the broken working data path.

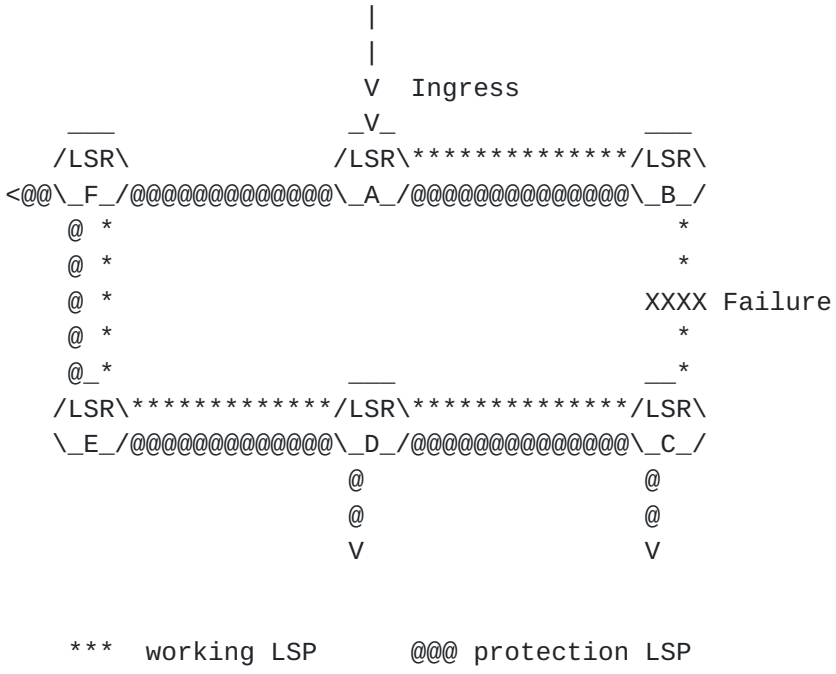

Figure 7: P2MP ROM Wrapping

Using this mechanism, there is a need to configure a particular protection LSP for each node on the working LSP. In the table below, "X's Backup" is the backup path activated by node X as a consequence of a failure affecting node Y (downstream node with respect to X) or link X-Y, and square brackets, in the path,indicate egress nodes.

Protected LSP: A->B->{C}->{D}->E->{F}

-- LINK/NODE PROTECTION --

A's Backup:	A->{F}->E->{D}->{C}
B's Backup:	B->A->{F}->E->{D}->{C}
C's Backup:	C->B->A->{F}->E->{D}
D's Backup:	D->C->B->A->{F}
E's Backup:	E->D->C->B->A->{F}

It should be noted that ROM-Wrapping is an LSP based protection mechanism, as opposed to the SPME based protection mechanisms that are presented in other sections of this draft. While this may seem to be limited in scope, the mechanism may be very efficient for many applications that are based on P2MP distribution schemes. While ROM-Wrapping can be applied to any network topology, it is particularly efficient for interconnected ring topologies.

3.1.1. Comparison of Wrapping and ROM-Wrapping

It is possible to compare the Wrapping and the ROM-Wrapping mechanisms in different aspects, and show some improvements offered by ROM-Wrapping.

When configuring the protection LSP for Wrapping it is necessary to configure for a specific failure: link protection or node protection. If the protection method is configured to protect node failures but the actual failure affects a link, this could result in failing to deliver traffic to the node, when it should be possible to.

ROM-Wrapping however does not have this limitation, because there is no distinction between node and link protection. Whether link B-C or node C fails, in either case the rerouting will attempt to reach C. If the failure is on the link, the traffic will be delivered to C, while if the failure is at node C, the traffic will be rerouted correctly until node D, and will be blocked at this point. However, all egress nodes up-to the failure will be able to deliver the traffic properly.

A second aspect is the number of hops needed to properly deliver the traffic. Referring to the example shown in Figure 7, where a failure is detected on link B-C, the following table lists the set of nodes traversed by the data in the protection:

Basic Wrapping:

A-B "Upstream" segment with respect to the failure	B-A-F-E-D-C backup path	{C}-{D}-E-{F} "Downstream" segment with respect to the failure

ROM Wrapping:

A-B "Upstream" segment with respect to the failure	B-A-{F}-E-{D}-{C} backup path	..

Comparing the two lists of nodes, it is possible to see that in this particular case the number of hops crossed using the simple Wrapping is significantly higher than the number of hops crossed by the traffic when ROM-Wrapping is used. Generally, the number of hops for basic Wrapping is always higher or at least equal compared to ROM-Wrapping. This implies a certain waste of bandwidth on all links that are crossed in both directions.

Considering the ring network previously seen, it is possible to do some bandwidth utilization considerations. The protected LSP is set up from A to F clockwise and an M Mbps bandwidth is reserved along the path. All the protection LSPs are pre-provisioned counterclockwise, each of them may also have reserved bandwidth M. These LSPs share the same bandwidth in a SE (Shared Explicit) [RFC2205] style.

The bandwidth reserved counterclockwise is not used when the protected LSP is properly working and could, in theory, be used for extra traffic [RFC4427]. However, it should be noted that [RFC5654] does not require support of such extra traffic.

The two recovery mechanism require different protection bandwidths. In the case of Wrapping, the bandwidth used is M in both directions of many of the links. While in case of ROM-Wrapping, only the links from the ingress node to the node performing the actual wrapping utilize M bandwidth in both directions, while all other links utilize M bandwidth only in the counterclockwise direction.

Consider the case of a failure detected on link B-C as shown in Figure 7. The following table lists the bandwidth utilization on each link (in units equal to M), for each recovery mechanism and for each direction (CW=clockwise, CCW=counterclockwise).

	Wrapping	ROM-Wrapping
Link A-B	CW+CCW	CW+CCW
Link A-F	CCW	CCW
Link F-E	CW+CCW	CCW
Link E-D	CW+CCW	CCW
Link D-C	CW+CCW	CCW

3.1.2. Multiple Failures Comparison

A further comparison between Wrapping and ROM-Wrapping can be done with respect to their ability to react to multiple failures. The wrapping recovery mechanism does not have the ability to recover from multiple failures on a ring network, while ROM-Wrapping is able to recover, from some multiple failures.

Consider, for example, a double link failure affecting links B-C and C-D shown in Figure 7. The Wrapping mechanism is not able to recover from the failure because B, upon detecting the failure, has no alternative paths to reach C. The whole P2MP traffic is lost. The ROM-Wrapping mechanism is able to partially recover from the failure, because the backup P2MP LSP to node F and node D is correctly set up and continues delivering traffic.

3.2. Steering for P2MP paths

When protecting P2MP traffic that uses an MPLS-TP ring as its branching point, i.e. it enters the ring at a head-end node and exits the ring at multiple nodes, we can employ a steering mechanism based on 1+1 linear protection [RFC6372]. We can configure two P2MP unidirectional SPME from each node on the ring that traverse the ring in both directions. These SPME will be configured with an egress at each ring node. In order to be able to properly direct the LSP traffic to the proper egress point for that particular LSP, we need to employ context labeling as defined in [RFC5331]. The method for using these labels is expanded upon in section 3.2.1.

For every LSP that enters the ring at a given node the traffic will be sent through both of these SPME, each with its own context label and the context-specific label for the particular LSP. The egress nodes should select the traffic that is arriving on the working SPME. When a failure condition is identified, the egress nodes should select the traffic from whichever of the two SPME whose traffic arrives at that node, i.e. since one of the two (presumably the working SPME) will be blocked by the failure. In this way, all egress nodes are able to receive the data traffic. While each node

detects that there is connectivity from the ingress point, it continues to select the data that is coming from the working SPME. If a particular node stops receiving the connectivity messages from the working SPME, it identifies that it must select to read the data packets from the protection SPME.

3.2.1. Context labels

Figure 8 shows the two unidirectional P2MP SPME that are configured from LSR-A with egress points at all of the nodes on the ring. The clockwise SPME (i.e. A-B-C-D-E-F) is configured as the working SPME, that will aggregate all traffic for P2MP LSPs that enter the ring at LSR-A and must be sent out of the ring at any subset of the ring nodes. The counter-clockwise SPME (i.e. A-F-E-D-C-B) is configured as the protection SPME.

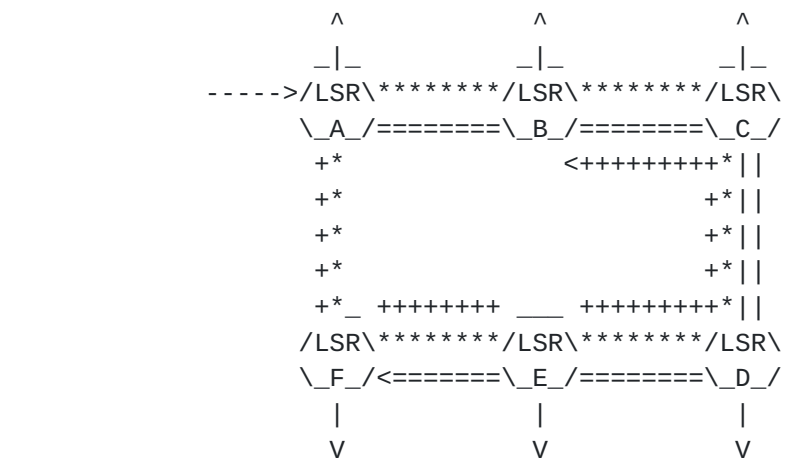

Figure 8: P2MP SPMEs

[RFC5331] defines the concept of context labels. A context-identifying label defines a context label space that is used to interpret the context-specific labels (found directly below the context- identifying label) for a specific tunnel. The SPME label is a context- identifying label. This means that at each hop the node that receives the SPME label uses it to point not directly to a forwarding table, but to a Label Information Base (LIB). As a node receives an SPME label it examines it, discovers that it is a context label, pops off the SPME label, and looks up the next label down in the stack in the LIB indicated by the context label.

The label below this context-identifying label should be used by the forwarding function of the node to decide the actions taken for this

packet. In MPLS-TP protection of ring topologies there are two context LIBs. One is the context LIB for the working SPME and the other is the context LIB for the P-SPME. All context LIBs have a behavior defined for the end-to-end LSP label but the behavior at each node may be different in the context of each SPME.

For example, using the ring that is shown in Figure 8, if the working SPME is configured to have a context-identifying label of CW at each node on the ring and the protection SPME is configured to have a context-identifying label of CP at each node. For the specific LSP we will designate the context-specific label used on the working SPME as WL(x-y) to be the label used as node-x to forward the packet to node-y. Similarly, for the context-specific labels on the protection SPME would be designated PL(x-y). An explicit example of label values appears in the next sub-section.

Applying 1+1 linear protection, as outlined above, for a P2MP LSP that enters the ring at LSR-A and has egress points from the ring at LSR-C and LSR-E using the two SPME shown in Figure 8 then a packet that arrives at LSR-A with a label stack [LI+S] will be forwarded on the working SPME with a label stack [CW | WL(A-B)]. The packet should then be forwarded to LSR-C arriving with a label [CW | WL(B-C)], where WL(B-C) should instruct the forwarding function to egress the packet with [LE(C)] and forward a copy to LSR-D with label stack [CW | WL(C-D)].

If a fault condition is detected, for example on the link C-D, then the nodes that are beyond the fault point, in this example nodes LSR-D, LSR-E, and LSR-F, will cease to receive the data packets from the clockwise (working) SPME. These LSR should then begin to switch their "selector bridge" and accept the data packets from the protection (counter-clockwise) SPME. At the ingress point, LSR-A, all data packets will have been transmitted on both the working SPME and the protection SPME. Continuing the example, LSR-A will transmit one copy of the data to LSR-B with stack [CW | WL(A-B)] and one copy to LSR-F with stack [CP | PL(A-F)]. The packet will arrive at LSR-C from the working SPME and egress from the ring. LSR-E will receive the packet from the protection SPME with stack [CP | PL(F-E)] and the context-sensitive label PL(F-E) will instruct the forwarding function to send a copy out of the ring with label LE(E) and a second copy to LSR-D with stack [CP | PL(E-D)]. In this way each of the egress points receives the packet from the SPME that is available at that point.

This architecture has the added advantages that there is no need for the ingress node to identify the existence of the mis-connectivity, and there is no need for a return path from the egress points to the ingress.

3.2.2. Walkthrough using context labels

In order to better demonstrate the use of the context labels we present a walkthrough of an example application of the P2MP protection presented in this section. Referring to Figure 9, there is a P2MP LSP that traverses the ring, entering the ring at LSR-B and branching off at LSR-D, LSR-E, and LSR-H and does not continue beyond LSR-H. For purposes of protection two P2MP unidirectional SPME are configured on the ring starting from LSR-B. One of the SPME, the working SPME, is configured with egress points at each of the LSR - C, D, E, F, G, H, J, K, A. The second SPME, the protection SPME, is configured with egress points at each of the LSR - A, K, J, H, G, F, E, D, C.

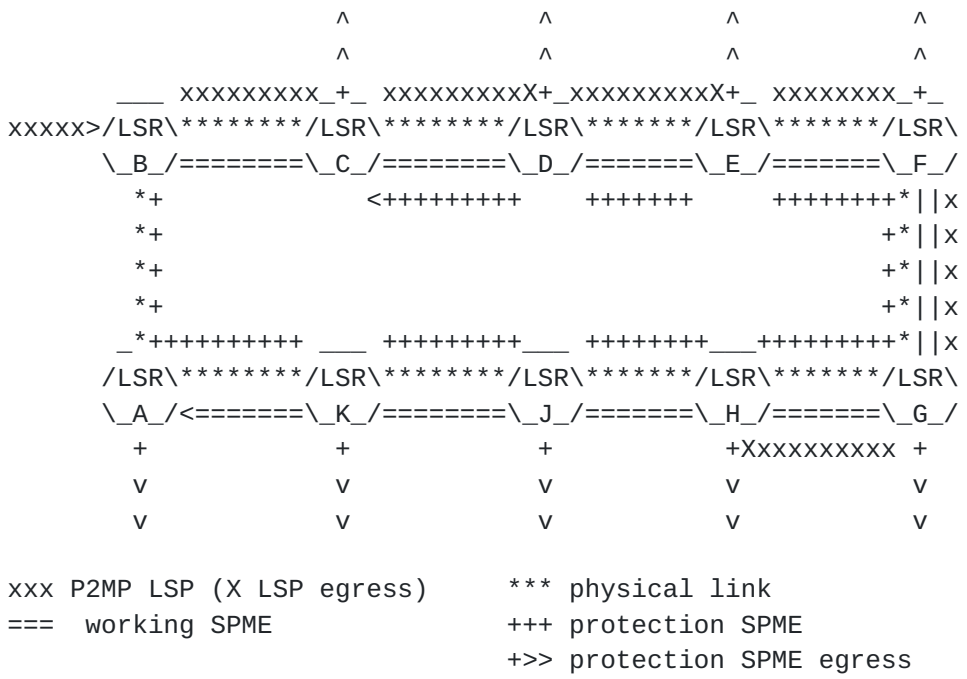

Figure 9: P2MP SPMEs

For this example we suppose that the LSP traffic enters the ring at LSR-B with the label stack [99], leaves the ring at LSR-D with stack [199], at LSR-E with stack [299], and LSR-H with stack [399].

While it is possible for the context-identifying label for the SPME be configured as a different value at each LSR, for the sake of this example we will suppose a configuration of 200 as the context-identifying label for the working SPME at each of the LSR in the ring, and 400 as the context-identifying label for the protection SPME at each LSR.

For the specific connected LSP we configure the following context-specific labels for each context:

node	W-context(200)	P-context(400)
A	65 {drop packet}	165 {fwrd w/[400|190]}
C	90 {fwrd w/[200|80]}	190 {drop packet}
D	80 {fwrd w/[200|75] + egress w/[199]}	180 {egress w/[199]}
E	75 {fwrd w/[200|65] + egress w/[299]}	175 {fwrd w/[400|180] + egress w/[299]}
F	65 {fwrd w/[200|55]}	165 {fwrd w/[400|175]}
G	55 {fwrd w/[200|45]}	155 {fwrd w/[400|165]}
H	45 {egress w/[399]}	145 {fwrd w/[400|155] + egress w/[399]}
J	65 {drop packet}	165 {fwrd w/[400|145]}
K	65 {drop packet}	190 {fwrd w/[400|165]}

When a packet arrives on the LSP to LSR-B with stack [99], the forwarding function determines that it is necessary to forward the packet to both the working SPME with stack [200|90] and the protection SPME with stack [400|165]. Each LSR on the SPME will identify the top label, i.e. 200 or 400, to be the context-identifying label and use the next label in the stack to select the forwarding action from the specific context table.

Therefore, at LSR-C the packet on the working SPME will arrive with stack [200|90] and the 200 will point to the table in the middle column above. After popping the 200 the next label, i.e. 90, will select the forwarding action "fwrd w/[200|80]" and the packet will be forwarded to LSR-D with stack [200|80]. In this manner, the packet will be forwarded along both SPME according to the configured behavior in the context tables. However, the egress points at LSR D, E, & H, will all be configured with a selector bridge to only use the input from the working SPME. If any of these egress points identify that there is a connection fault on the working SPME, then the selector bridge will cause the LSR to read the input from the protection SPME.

4. Coordination protocol

The Survivability Framework [RFC6372] indicates that there is a need to coordinate protection switching between the end-points of a protected bidirectional domain. The coordination is necessary for particular cases, in order to maintain the co-routed nature of the

bidirectional transport path. The particular cases where this becomes necessary include cases of unidirectional fault detection and use of operator commands.

By using the same mechanisms defined in [RFC6378], for linear protection, to apply for protection of a single ring topology we are able to gain a consistent solution for this coordination between the end-points of the protection domain. The Protection State Coordination Protocol that is specified in [RFC6378] provides coverage for all the coordination cases, including support for operator commands, e.g. Forced-Switch.

5. Conclusions and Recommendations

Ring topologies are prevalent in traditional transport networks and will continue to be used for various reasons. Protection for transport paths that traverse a ring within an MPLS network can be provided by applying an appropriate instance of linear protection, as defined in [RFC6372]. This document has shown that for each of the traditional ring protection architectures there is an application of linear protection that provides efficient coverage, based on the use of the Sub-Path Maintenance Entity (SPME), defined in [RFC5921] and [RFC6371]. For example,

- P2P Steering - Configuration of two SPME, from ring ingress to ring egress, and 1:1 linear protection

- P2P Wrapping for link protection - Configuration of two SPME, one for the protected link and the second using the long route between the two neighboring nodes, and 1:1 linear protection.

- P2P Wrapping for node protection - Configuration of two SPME, one between the two neighbors of the protected node and the second between these two nodes on the long route, and 1:1 linear protection.

- P2MP Wrapping - it is possible to optimize the performance of the wrapping by configuring the proper protection path to egress the data at the proper branching nodes.

- P2MP Steering - by combining 1+1 linear protection and configuration of the SPME based on context-sensitive labeling of the protection path.

It has been shown that this set of protection architecture and mechanisms are optimized based on the criteria defined in [RFC5654] for justification of designing a specific protection mechanism for a

ring topology.

Protection of traffic over a ring topology based on the Steering architecture using basic 1:1 linear protection is a very efficient implementation for sections of a P2P transport path that traverses a ring. Steering should be the preferred mechanism for P2P protection in a ring topology since it reduces the extra bandwidth required when traffic doubles through wrapped protection, and the ability to protect both against link and node failures without complicating the fault detection or the need to configure multiple protection paths. While this is true, the possiblity remains to support either mechanism while depending upon the OAM functionality [outlined in [RFC6371] and specified in various documents] and the coordination protocol specified for linear protection in [RFC6378].

6. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

7. Security Considerations

This document does not add any security risks to the network. Any security considerations are defined in [RFC6378] and their applicability to the information contained in this document follow naturally from the applicability of the mechanism defined in that document.

8. Acknowledgements

The authors would like to acknowledge the strong contributions from all the people commenting on this draft and making suggestions for improvements.

9. References

9.1. Normative References

[RFC6378] Weingarten, Y., Bryant, S., Osborne, E., Sprecher, N., and A. Fulignoli, "MPLS-TP Linear Protection", RFC 6378, October 2011.

9.2. Informative References

[RFC4090] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

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

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

[RFC5921] Bocci, M., Bryant, S., Frost, D., Levrau, L., and L. Berger, "MPLS-TP Framework", RFC 5921, July 2010.

[RFC6371] Busi, I. and D. Allan, "MPLS-TP OAM Framework", RFC 6371, Sept 2011.

[RFC6372] Sprecher, N. and A. Farrel, "MPLS-TP Survivability Framework", RFC 6372, Sept 2011.

[RFC2205] Braden, R., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) - Functional Specifications", RFC 2205, September 1997.

[RFC4427] Mannie, E. and D. Papadimitriou, "Recovery (Protection and Restoration) Terminology for GMPLS", RFC 4427, March 2006.

[G.841] ITU, "Types and characteristics of SDH network protection architectures", ITU-T G.841, October 1998.

Authors' Addresses

Yaacov Weingarten
34 Hagefen St.
Karnei Shomron, 4485500
Israel

Phone:
Email: wyaacov@gmail.com

Stewart Bryant
Cisco
United Kingdom

Email: stbryant@cisco.com

Danielle Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova, Sestri Ponente
Italy

Email: daniele.ceccarelli@ericsson.com

Diego Caviglia
Ericsson
Via A. Negrone 1/A
Genova, Sestri Ponente
Italy

Email: diego.caviglia@ericsson.com

Francesco Fondelli
Ericsson
Via A. Negrone 1/A
Genova, Sestri Ponente
Italy

Email: francesco.fondelli@ericsson.com

Marco Corsi
Altran
Via A. Negrone 1/A
Genova, Sestri Ponente
Italy

Email: corsi.marco@gmail.com

Bo Wu
ZTE Corporation
4F,RD Building 2,Zijinghua Road
Nanjing, Yuhuatai District
P.R.China

Email: wu.bo@zte.com.cn

Xuehui Dai
ZTE Corporation
4F,RD Building 2,Zijinghua Road
Nanjing, Yuhuatai District
P.R.China

Email: dai.xuehui@zte.com.cn