Applicability of MPLS Transport Profile for Ring Topologies

Abstract

   This document presents an applicability of existing MPLS protection
   mechanisms, both local and end-to-end, to the MPLS Transport Profile
   (MPLS-TP) in ring topologies. This document does not propose any new
   mechanisms or protocols. 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 discusses how most of the requirements are met by
   applying linear protection as defined in RFC 6378 in a ring topology.

Copyright Notice

Table of Contents

1. Introduction

   The MPLS Transport Profile (MPLS-TP) has been standardized as part of
   a joint effort between the Internet Engineering Task Force (IETF) and
   the International Telecommunications Union Telecommunications
   Standardization Sector (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 subnetworks 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.
   The requirements state that specific protection mechanisms applying
   to ring topologies may be developed if these allow the network to
   minimize:

   o the number of OAM entities needed to trigger the protection

   o the number of elements of recovery needed

   o the number of labels required

   o the number of control- and management-plane transactions during a
      maintenance operation

   o the 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 of the following cases.

   1. A point-to-point transport path that originates at a ring node or
       enters an MPLS-TP-capable ring at a single ingress node, and
       exits the ring at a single egress node, and possibly continues
       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 it 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 it should be detected by the two neighboring ring
       nodes.

   3. An operator command is issued to a specific ring node; it either
       changes the operational state of a node or a link or explicitly
       triggers a protection action. 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, and 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 ingress node of the ring 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:

o MPLS-TP framework [RFC5921]

o MPLS-TP OAM framework [RFC6371]

o 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 (LSRs) 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 node of 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". "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 Pxi(y) in the stack denotes the label that LSR-x would
    use to transport the packet to LSR-y over the SPME whose index is
    i.

   For example:

   o The label stack [LI] denotes the label stack received at the
      ingress node of the ring. There may be additional labels after
      LI, e.g., a PW label; however, this is irrelevant to the
      discussion of the protection scenario.

   o [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, and the second
      label is the label that would be used by the egress LSR to
      continue to transmit the packet on the original LSP.

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

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

   There are two protection architecture mechanisms -- "Wrapping" and
   "Steering" -- that have historically been applied to ring topologies,
   based on Synchronous Digital Hierarchy (SDH) specifications [G.841],
   and have been proposed in various forums to perform recovery of a
   topological ring network. The following subsections 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 to
   transmit the data along the path, there is a need to "wrap" all data
   traffic around the ring, on an alternate data path, until the arrives
   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 making a local decision, the far-side
node can wrap the data back onto the normal working path until the
egress from the ring segment.

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

```
                 ___ ######## ___ ######## ___
          ======>/LSR\********/LSR\***XX***/LSR\
             \_B_/@@@@@@@@\_A_/         \_F_/
               *@                        #*@
               *@                        #*@
               *@                        #*@
              _*@            ___         #*@
             /LSR\********/LSR\********/LSR\======>
             \_C_/@@@@@@@@\_D_/@@@@@@@@\_E_/


            ==> connected LSP  *** physical link
            ###  working path   @@@ wrapped data path
```
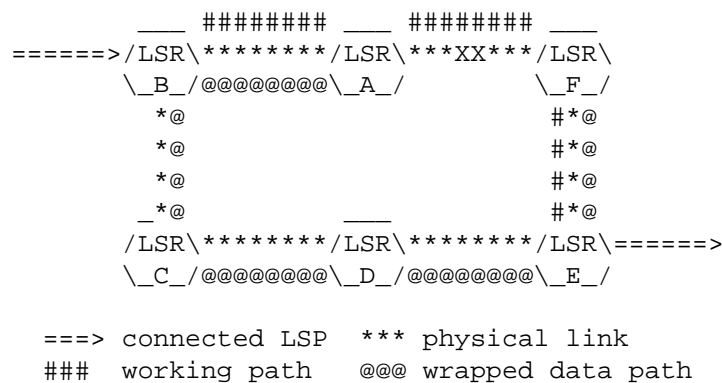
              Figure 1: Wrapping Protection for P2P Path

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 LSRs 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:

o Detection of mis-connectivity or loss of continuity 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.

o 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 an
   LSP, then, for a ring with N nodes, there is a need for O(2N)
   alternate paths.

o 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, and later it
   is transmitted E->F and F->E. This means that there is additional
   bandwidth needed for this protection.

o 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.

o 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
   zero resources and depending on the Network Management System
   (NMS) to allocate the proper resources around the ring, based on
   actual traffic usage.

o 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
   node of 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).

```
                    ___         ___         ___
           ======>/LSR\*******/LSR\*******/LSR\======>
                  \_B_/#######\_A_/#######\_F_/
                   *@                       @*
                   *@                       @*
                   *@                       @*
                  _*@           ___         @*_
                  /LSR\*******/LSR\*******/LSR\
                  \_C_/@@@@@@@\_D_/@@@@@@@\_E_/

              ===> connected LSP     *** physical link
              ###  working path      @@@ protection path
```

              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. This requires that the ingress node be informed of the need
to switch over to the protection path, and also that the ingress and
egress nodes coordinate the switchover. 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
endpoints of the protected domain.

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

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

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

o 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 types of transport
   path and protection that are being implemented. This will be
   detailed in the following subsections.

   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 depending 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 switchover 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] and 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 LSPs 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 SPMEs through the ring (the first SPME
   traverses B-A-F, and the second SPME traverses B-C-D-E-F).

                        ___   ___   ___
               =====>/LSR********/LSR********/LSR\======>
                     _B_/########_A_/########_F_/
                      @   @
                      @   @
                      @   @
                     _*@   ___   @*_
                     /LSR********/LSR********/LSR\
                     _C_/@@@@@@@@_D_/@@@@@@@@_E_/

                 ===> connected LSP   *** physical link
                 ### primary SPME   @@@ secondary SPME
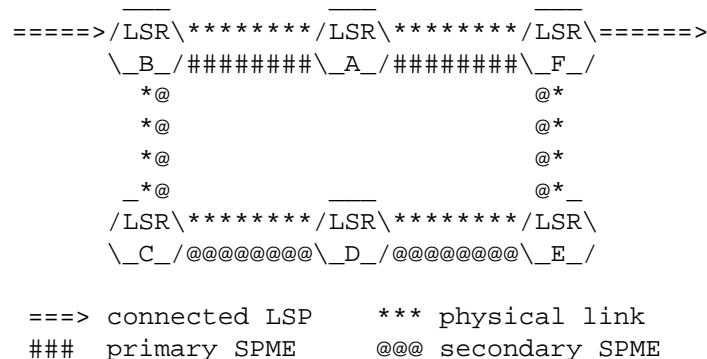
                      Figure 3: An MPLS-TP Ring

   This protection mechanism is identical to the 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 the OAM
   indication, another 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 SPMEs -- the
first is a SPME between the two LSRs that are connected by the link,
and the second SPME is between those same two LSRs but traverses the
entire ring (except the link that connects the LSRs). In Figure 4,
we show the primary SPME that connects LSR-A and LSR-F over a segment
connection, and the secondary SPME that connects these same LSRs by
traversing the ring in the opposite direction.

```
               ___         ___         ___
              /LSR\********/LSR\********/LSR\
              \_B_/@@@@@@@@\_A_/########\_F_/
                *@                        *@
                *@                        *@
                *@                        *@
               _*@         ___          _*@
              /LSR\********/LSR\********/LSR\
              \_C_/@@@@@@@@\_D_/@@@@@@@@\_E_/


         *** physical link
         ### primary SPME      @@@ secondary SPME
```
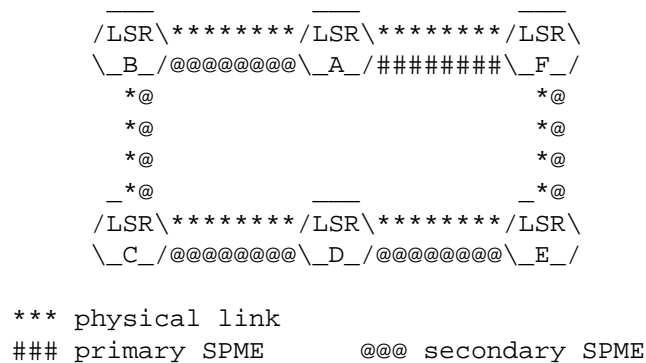
Figure 4: Segment SPMEs

By applying OAM monitoring of these two SPMEs (at each LSR), it is
possible to effect 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.,
    the 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., swaps the label
    for the LSP, to be acceptable to the SPME egress, and pushes 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 pushes 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 pops the SPME label, process
    the LSP label, and forwards 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 subsection. The difference
   would be in the SPMEs that are used. For node protection, the
   primary SPME would be configured between the two LSRs that are
   connected to the node that is being protected (see the 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 the secondary SPME in Figure 5).
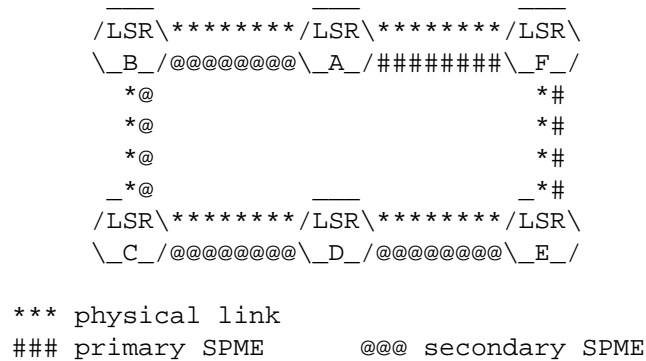
```
                  ___          ___          ___
                 /LSR\********/LSR\********/LSR\
                 \_B_/@@@@@@@@\_A_/########\_F_/
                  *@                        *#
                  *@                        *#
                  *@                        *#
                 _*@          ___          _*#
                 /LSR\********/LSR\********/LSR\
                 \_C_/@@@@@@@@\_D_/@@@@@@@@\_E_/


          *** physical link
          ### primary SPME       @@@ secondary SPME


              Figure 5: Node-Protection SPMEs
```

The protection mechanism would work similarly -- it would be based on
1:1 linear protection [RFC6372] and be triggered by OAM functions on
both SPMEs. It would wrap 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 against 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 SPMEs 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 SPMEs
configured:

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

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

o 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 the traffic to the path between LSR-A and
   LSR-F on the long circuit of the ring.

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

```
                  ___ ++++++++ ___                   ___
                 /LSR\********/LSR\********/LSR\
                 \_B_/@@@@@@@@\_A_/########\_F_/
                 $+*@                        +*$
                 $+*@                        +*$
                 $+*@                        +*$
                 $+*@ ++++++++ ___ ++++++++ +*$
                 /LSR\********/LSR\********/LSR\
                 \_C_/@@@@@@@@\_D_/@@@@@@@@\_E_/
                    $$$$$$$$      $$$$$$$$

          *** physical link
          ### primary SPME          @@@ secondary node#1 SPME
          $$$ secondary node#2 SPME  +++ secondary segment SPME
```
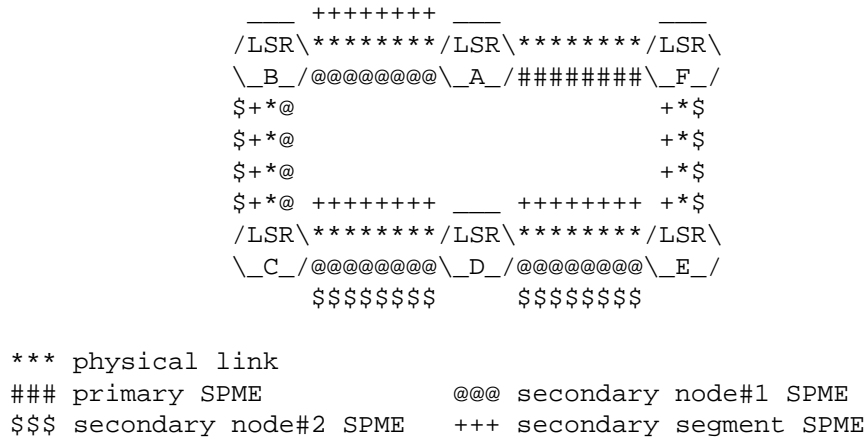
            Figure 6: SPMEs for Protecting Segments and Node

2.4. Analysis of P2P Protection

   Analyzing steering SPME protection (Section 2.3.1) and wrapping based
   on SPME (Sections 2.3.2 or 2.3.3), we can make the following
   observations (based on a ring with N nodes, where N is not more than
   16):

   o Number of SPMEs that need to be configured

         For steering: O(2N^2). There are two SPMEs from each ingress
         LSR to each of the other nodes in the ring.

         For wrapping: O(2N). (However, the operator must decide a
         priori whether to protect for link failures or node failures at
         each point.)

   o Number of OAM sessions at each node

         For steering: O(2N)

         For wrapping: 3

   o Bandwidth requirements

         For 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.

   o Special considerations

         For steering: latency of OAM detection of fault condition by
         ingress MEP. (Using alarm reporting could optimize over using
         CC-V only.)

         For wrapping: each node must decide a priori whether it is
         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 violate 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 SPMEs involved in this
   protection could be reduced by eliminating each SPME between a pair
   of LSRs that is 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 in a ring topology is possible as
   demonstrated, there are certain limitations in addressing some of the
   required behavior. The limitations include:

   o the need to configure a priori whether link or node protection
      will be provided

   o the higher number of SPMEs that need to be defined

   o the difficulty in addressing cases of multiple failures in the
      ring

   Application of linear protection, based on the use of SPMEs within
   the ring, to implement a steering methodology to protect a ring
   topology is rather straightforward, 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 P2P paths that traverse a ring topology.

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 subsections will present methods to address the protection
   of the ring-based sections of these LSPs.

3.1. Wrapping for P2MP LSPs

   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 a protection P2MP LSP, relative to
   each node that is considered a failure risk. The protection P2MP LSP
   will be routed between the failure risk node's upstream neighbor to
   all of the egress nodes (for the particular LSP) that are downstream
   of the failure risk node.

   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, and C) complementary to the broken working
   data path.

```
                                |
                                |
                                V  Ingress
          ___                  _V_                    ___
         /LSR\               /LSR\*************/LSR\
      <@@\_F_/@@@@@@@@@@@@@\_A__/@@@@@@@@@@@@@\_B_/
         @  *                                        *
         @  *                                        *
         @  *                            XXXX Failure
         @  *                                        *
         @_*              ___                      __*
         /LSR\*************/LSR\*************/LSR\
         \_E_/@@@@@@@@@@@@@\_D__/@@@@@@@@@@@@@\_C_/
                           @                     @
                           @                     @
                           V                     V


              ***  working LSP      @@@ protection LSP
```
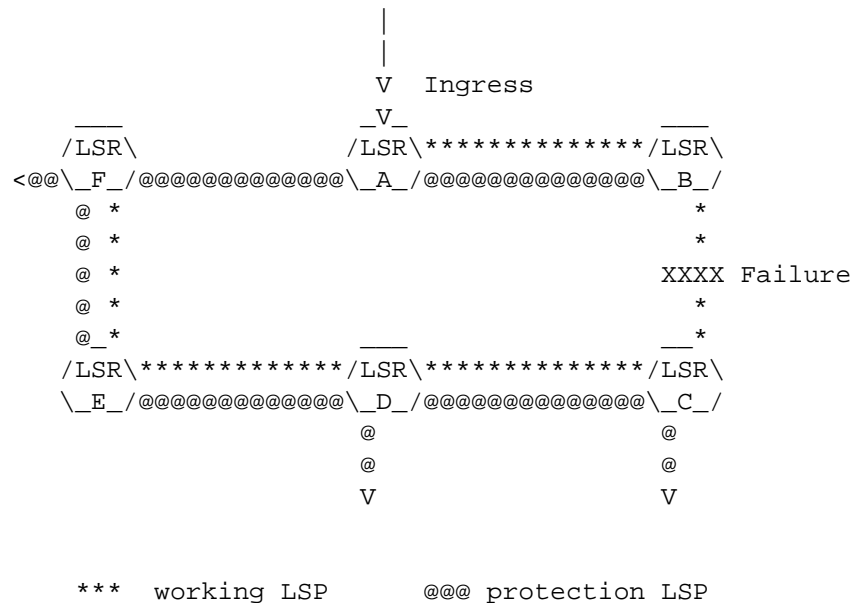
Figure 7: P2MP ROM-Wrapping

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. (Note: Braces 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 document. 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 various 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 against 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
   do so.

   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, the rerouting will attempt to reach C. If the failure
   is on the link, the traffic will be delivered to C; 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   B-A-F-E-D-C   {C}-{D}-E-{F}
   "Upstream" segment   backup path   "Downstream" segment
   with respect to the   with respect to the
   failure   failure

                          ROM-Wrapping:

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

   Comparing the two lists of nodes, it is possible to see that in this
   particular case the number of hops crossed when basic wrapping is
   used 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 greater than or equal to that for ROM-
   Wrapping. This implies a certain waste of bandwidth on all links
   that are crossed in both directions.

Considering the ring network in Figure 7, it is possible to consider
the bandwidth utilization. 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) style, as described in [RFC2205].

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

The two recovery mechanisms require different protection bandwidths.
In the case of wrapping, the bandwidth used is M in both directions
on many of the links. While in the 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 of 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. All the P2MP traffic is lost. The

ROM-Wrapping mechanism is able to partially recover from the failure,
because the backup P2MP LSP to F and 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., the traffic 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 SPMEs from each node on the ring; they
traverse the ring in both directions. These SPMEs will be configured
with an egress at each ring node. In order to be able to 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 SPMEs, 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 SPMEs 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 node of the ring,
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 SPMEs 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
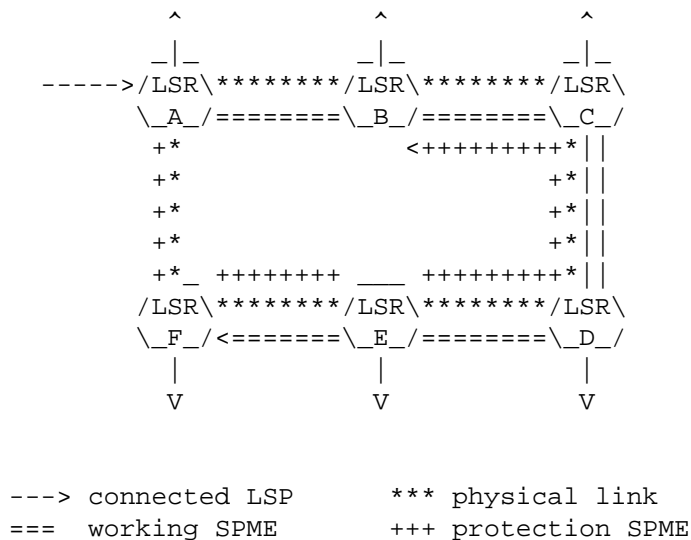as the protection SPME.

```
                        ^               ^               ^
                       _|_             _|_             _|_
              ----->/LSR\********/LSR\********/LSR\
                    \_A_/========\_B_/=======\_C_/
                     +*                  <+++++++++*||
                     +*                         +*||
                     +*                         +*||
                     +*                         +*||
                     +*_ ++++++++ ___ +++++++++*||
                    /LSR\********/LSR\********/LSR\
                    \_F_/<=======\_E_/=======\_D_/
                     |               |               |
                     V               V               V


              ---> connected LSP      *** physical link
              ===  working SPME       +++ 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 to take 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 protection 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, 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), where it's the label used as node-x forwards the packet
to node-y. Similarly, a context-specific label on the protection
SPME would be designated PL(x-y). An explicit example of label
values appears in the next subsection.

Assume we are 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 SPMEs shown in Figure 8. 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. Each of these LSRs should then begin
to switch its "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. Walk-Through Using Context Labels

In order to better demonstrate the use of the context labels, we
present a walk-through 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 it does not continue
beyond LSR-H. For purposes of protection, two P2MP unidirectional
SPMEs are configured on the ring starting from LSR-B. One of the
SPMEs, the working SPME, is configured with egress points at each of
the LSRs -- C, D, E, F, G, H, J, K, A. The second SPME, the
protection SPME, is configured with egress points at each of the LSRs
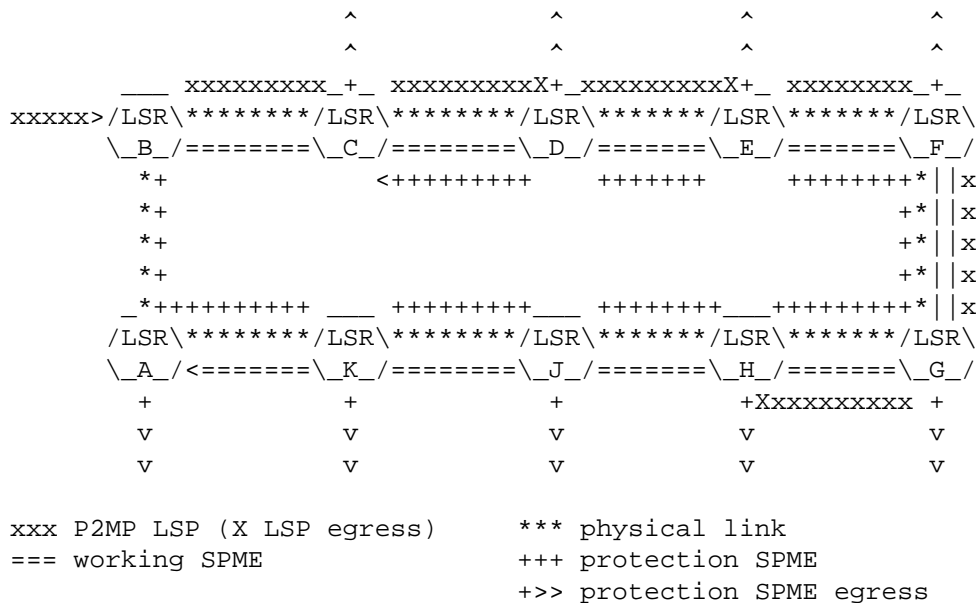-- A, K, J, H, G, F, E, D, C.

```
                       ^              ^            ^             ^
                       ^              ^            ^             ^
          ___ xxxxxxxxx_+_ xxxxxxxxxX+_xxxxxxxxxX+_ xxxxxxxxx_+_
     xxxxx>/LSR\*******/LSR\*******/LSR\******/LSR\*******/LSR\
          \_B_/=======\_C_/=======\_D_/======\_E_/=======\_F_/
           *+              <+++++++++    +++++++    ++++++++*||x
           *+                                            +*||x
           *+                                            +*||x
           *+                                            +*||x
          _*+++++++++ ___ +++++++++___ +++++++___+++++++++*||x
          /LSR\*******/LSR\*******/LSR\******/LSR\*******/LSR\
          \_A_/<======\_K_/=======\_J_/======\_H_/=======\_G_/
           +             +             +          +Xxxxxxxxxx +
           v             v             v          v          v
           v             v             v          v          v

          xxx P2MP LSP (X LSP egress)    *** physical link
          === working SPME              +++ protection SPME
                                        +>> protection SPME egress
```

                        Figure 9: P2MP SPMEs

   For this example, we suppose that the LSP traffic enters the ring at
   LSR-B with the label stack [99], and leaves the ring:

   o at LSR-D with stack [199]

   o at LSR-E with stack [299]

   o at LSR-H with stack [399]

   While it is possible for the context-identifying label for the SPME
   to 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 LSRs 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:

```
+------+----------------------------+----------------------------+
| node | W-context(200)             | P-context(400)             |
+------+----------------------------+----------------------------+
|   A  | 65 {drop packet}           | 165 {fwd w/ [400 | 190]}   |
|   C  | 90 {fwd w/ [200 | 80]}     | 190 {drop packet}          |
|   D  | 80 {fwd w/ [200 | 75] +    | 180 {egress w/ [199]}      |
|      | egress w/ [199]}           |                            |
|   E  | 75 {fwd w/ [200 | 65] +    | 175 {fwd w/ [400 | 180] +  |
|      | egress w/ [299]}           | egress w/ [299]}           |
|   F  | 65 {fwd w/ [200 | 55]}     | 165 {fwd w/ [400 | 175]}   |
|   G  | 55 {fwd w/ [200 | 45]}     | 155 {fwd w/ [400 | 165]}   |
|   H  | 45 {egress w/ [399]}       | 145 {fwd w/ [400 | 155] +  |
|      |                            | egress w/ [399]}           |
|   J  | 65 {drop packet}           | 165 {fwd w/ [400 | 145]}   |
|   K  | 65 {drop packet}           | 190 {fwd 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 middle column of the
table above. After popping the 200, the next label, i.e., 90, will
select the forwarding action "fwd 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 SPMEs according to the configured
behavior in the context tables. However, the egress points at LSR-D,
LSR-E, and LSR-H will each be configured with a selector bridge so
they will use only the input from the working SPME. If any of these
egress points identifies 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 endpoints 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 when unidirectional fault detection or
   operator commands are used.

   By using the same mechanisms defined in [RFC6378] for linear
   protection to protect a single ring topology, we are able to gain a
   consistent solution for this coordination between the endpoints 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:

   o P2P steering - Configuration of two SPMEs, from the ingress node
      of the ring to the egress node of the ring, and 1:1 linear
      protection.

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

   o P2P wrapping for node protection - Configuration of two SPMEs, 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.

   o 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.

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

   This document shows that use of the protection architecture and
   mechanisms suggested provides the optimizations needed to justify
   ring-specific protection as defined in [RFC5654].

   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 it provides the
   ability to protect both against link and node failures without
   complicating the fault detection or requiring that multiple
   protection paths be configured. While this is true, it's possible to
   support either wrapping or steering 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. 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 follows
   naturally from the applicability of the mechanism defined in that
   document.

7. References

7.1. Normative References

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

7.2. Informative References

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

   [RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S.
              Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1
              Functional Specification", RFC 2205, September 1997.

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

   [RFC4427] Mannie, E. and D. Papadimitriou, "Recovery (Protection and
              Restoration) Terminology for Generalized Multi-Protocol
              Label Switching (GMPLS)", RFC 4427, March 2006.

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

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

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

   [RFC6371] Busi, I. and D. Allan, "Operations, Administration, and
              Maintenance Framework for MPLS-Based Transport Networks",
              RFC 6371, September 2011.

   [RFC6372] Sprecher, N. and A. Farrel, "MPLS Transport Profile
              (MPLS-TP) Survivability Framework", RFC 6372,
              September 2011.

Appendix A. Acknowledgements

   The authors would like to acknowledge the strong contributions from
   all the people who commented on this document and made suggestions
   for improvements.

Appendix B. Contributors

   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)

Authors' Addresses

   Yaacov Weingarten
   34 Hagefen St.
   Karnei Shomron,   4485500
   Israel

   Phone:
   EMail: wyaacov@gmail.com

   Stewart Bryant
   Cisco Systems
   10 New Square, Bedfont Lakes
   Feltham, Middlesex,
   TW18 8HA
   UK

   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

EMail: xuehuiwfsy@gmail.com