MPLS Working Group T. Cheung
Internet-Draft J. Ryoo
Intended status: Standards Track ETRI
Expires: September 13, 2012 Y. Weingarten
N. Sprecher
Nokia Siemens Networks
D. King
Old Dog Consulting
March 12, 2012

MPLS-TP Shared Mesh Protection
draft-cheung-mpls-tp-mesh-protection-05.txt

Abstract

This document describes a mechanism to address the requirement to support protection of Label Switched Paths (LSPs) in an MPLS Transport Profile (MPLS-TP) mesh topology. The shared mesh protection mechanism enables multiple protection paths within a shared mesh protection domain to share protection resources for the protection of working paths by coordinating protection switching operations according to the priority assigned to each end-to-end linear protection domain.

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 3
2. Conventions Used in this Document 5
  2.1. Acronyms 5
  2.2. Definitions and Terminology 5
3. Shared Mesh Protection Architecture 6
  3.1. Shared Mesh Protection Group 7
  3.2. Shared Start and End Nodes 8
  3.3. SMP protocol communication channel 10
  3.4. Network planning for SMP 10
  3.5. Preemption and race conditions 11
  3.6. SMP Overview 12
    3.6.1. LP Protocol extensions for shared protection 12
    3.6.2. Notifying protection switching event 13
    3.6.3. Requesting lockout of protection 13
    3.6.4. Resolving race conditions 14
4. Protocol 15
  4.1. PDU Format 15
  4.2. Message Transmission 17
5. Operation of Shared Mesh Protection 17
6. Manageability Considerations 20
7. IANA Considerations 21
8. Security Considerations 21
9. References 21
  9.1. Normative References 21
  9.2. Informative References 21
Authors' Addresses 22

1. Introduction

The MPLS Transport Profile (MPLS-TP) is a packet transport technology based on a profile of the MPLS and Pseudowires (PW) as described in [RFC3031], [RFC3985], and [RFC5085]. MPLS-TP is the application of MPLS to the construction of packet-switched paths that are analogous to traditional circuit-switched technologies. Requirements for MPLS-TP are specified in [RFC5654].

An important feature of a transport network is its survivability function and the ability to maintain or recover traffic following a network failure or attack. According to Requirement 56 of [RFC5654], MPLS-TP must provide protection and restoration mechanisms, and it must also be possible to protect 100% of the traffic on the protected path (Requirement 58).

1+1 and 1:1 linear protection meets these requirements by reserving the equivalent amount of network resources for the protection paths as is allocated to the normal traffic that is being protected. While those dedicated protection mechanisms provide very good protection capabilities, they are resource inefficient and will increase overall network resource consumption. Deploying 1+1 and 1:1 protection mechanisms for all services that require resiliency, dramatically increases network costs.

[RFC5654] also establishes that MPLS-TP should support shared protection (Requirement 68). 1:n end-to-end protection uses one protection path to protect n working paths between the same two endpoints. This improves overall network utilization, but the resource (bandwidth) allocated to a protection path is typically not sufficient to protect multiple simultaneous failures on different working paths. If multiple working paths require concurrent protection switching, the path with the highest priority should be protected as described in [RFC6372].

In 1+1 and 1:1 protection, the end nodes of the working path must be the same as those of the protection path. Similarly in 1:n protection all pairs of end nodes of the n working paths are the same, and the protection path must also have the same end nodes. In the event that the MPLS-TP network scales up, the number of Label Switched Paths (LSPs) having different end nodes will also increase. The network utilization benefit for sharing protection resources among multiple protected domains for such LSPs will increase accordingly.

Requirement 68 of [RFC5654] specifies that MPLS-TP should support 1:n shared mesh recovery, and Requirement 69 states that MPLS-TP must support sharing of protection resources. It may be possible that

some working paths are sufficiently disjoint and would be unlikely to be simultaneously affected by a single network failure. Typically, such a scenario is hard to track in real network environments where new services are often added and removed.

In mesh protection, network resources may be shared to provide protection for working paths that do not share the same end nodes at the edge of a protection domain. This type of protection can make very efficient use of network resources, but requires coordination of several segments in order to ensure that only a single traffic flow is switched to the protection resources at any time.

[RFC4428] defines two shared mesh recovery schemes named (1:1)^n and (M:N)^n. The (1:1)^n recovery scheme is a simple case of (M:N)^n recovery scheme. In (1:1)^n protection, n working paths are protected by n dedicated protection paths while sharing the same protection bandwidth. The protection bandwidth can be optimized to allow only one of the n working paths to be protected at any time. In this case, it achieves network utilization similar to 1:n protection.

It should be noted that the (1:1)^n protection scheme described in [RFC4428] differs with that defined in [G.808.1] in that the former allows each n pairs of working and protection paths to have different end nodes while the latter applies to the case where all pairs have the same end nodes.

This document defines a data-plane shared mesh protection mechanism based on the concept of the (1:1)^n recovery scheme described in [RFC4428] and a protocol for coordination of the shared protection resources. The actual protection switching is controlled by end-to-end linear protection, while the usage of the shared resources is based on the protection switching priority assigned to each pair of working and protection paths.

The shared mesh protection mechanism defined in this document utilizes the existing MPLS-TP linear protection switching mechanism, and assumes that the protection paths are established and ready to forward data prior to a failure. Upon detection of a failure on a working path, only the two end nodes of the failed working path exchange their linear protection protocol messages to switch data traffic. No explicit activation procedure to switch data traffic to the protection path is needed in the intermediate nodes along the protection path. However, the intermediate nodes that are part of the shared segments need to coordinate the resource allocation on the shared nodes and this coordination will be addressed by the protocol proposed in this document.

2. Conventions Used in this Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2.1. Acronyms

This draft uses the following acronyms:

G-ACh	Generic Associated Channel Header
LoP	Lockout of Protection
LP	Linear Protection
LSP	Label Switched Path
MIP	Maintenance Entity Group Intermediate Point
MPLS-TP	Transport Profile for MPLS
P2P	Point-to-point
P2MP	Point-to-multipoint
PS	Protection Switching
PW	Pseudowire
RA	Resource Allocation
SEN	Shared End Node
SMP	Shared Mesh Protection
SMPG	Shared Mesh Protection Group
SPME	Sub-Path Maintenance Entity
SRLG	Shared Risk Link Group
SSN	Shared Start Node

2.2. Definitions and Terminology

This document defines two protection domains as follows:

o End-to-end linear protection (LP) domain: A protection domain as defined in [RFC6372] for protecting a P2P or P2MP LSP. It consists of two or more endpoints at the boundary of the domain and a working path and a protection path between the end nodes. An end-to-end linear protection switching protocol runs within the domain.

o Shared mesh protection (SMP) domain: A protection domain for protecting a number of P2P or P2MP LSPs. It consists of a number of end-to-end linear protection domains. Each end-to-end linear protection domain shares protection resources with other domains. The shared protection resource may be a node, link, transport path segment or concatenated transport path segment. A shared mesh protection switching protocol runs within the domain.

In addition, we define the following:

o Protection segment: A protection segment is a portion of the end-to-end protection path. An end-to-end protection path can be broken down into separate protection segments. A protection segment may either be a shared protection segment or a dedicated protection segment. A shared protection segment is a segment that has a set of resources, e.g. link bandwidth that is reserved as shared amongst several end-to-end protection paths. A dedicated protection segment is a segment of the end-to-end protection path whose resources are reserved for use only by that protection path.

o Shared mesh protection group (SMPG): A protection group includes the pairs of working and protection paths, whose working paths do not belong to a single SRLG and whose protection paths share the resource of a single shared protection segment. Note that an LSP may belong to multiple SMPGs.

3. Shared Mesh Protection Architecture

The shared mesh protection domain shown in Figure 1 has two end-to-end linear protection domains. One consists of the two end nodes A and E and includes one working path, ABCDE, and one dedicated protection path APQRE. The second consists of end nodes V and Z and one working path, VWXYZ, and the dedicated protection path, VPQRZ. Those two LP domains include a shared protection segment PQR for their protection paths. This illustrates a simple configuration of shared mesh protection. Note that the two working paths, ABCDE and VWXYZ, do not share endpoints so they cannot make use of 1:n protection even though they also do not share any potential common points of failure.

It is possible to apply linear protection to each of these working paths individually. If there are no failures affecting either of the two working paths, the shared protection segment PQR carries no traffic (or only interruptible extra traffic). In the event of only one failure on a working path, the segment PQR carries traffic from the working path that detected the failure. Only in the event that there are failures detected on both of the working paths is there a conflict over the appropriate use of the shared protection segment PQR. It is important to note that there are two distinct LSPs (i.e. APQRE and VPQRZ) that are signaled over the shared protection segment and that although we refer to the singular segment, the traffic is actually being transported on separated transport paths.

Thus, it is possible for the network resources of segment PQR to be shared by the two protection paths. In this way, shared mesh protection can substantially reduce the amount of network resources that need to be reserved to provide protection of the multiple paths

within the same protection group.

```
      A----B----C----D----E
       \                 /
        \               /
         \             /
          P-----Q-----R
         /             \
        /               \
       /                 \
      V----W----X----Y----Z
```

Figure 1: A Shared Mesh Protection Topology

3.1. Shared Mesh Protection Group

In Figure 1, two working paths, ABCDE and VWXYZ and their corresponding protection paths, APQRE and VPQRZ, are considered a Shared Mesh Protection Group (SMPG). As pointed out above, each protection path belonging to a SMPG is an individual LSP, but it shares the resources with others in a segment. The resources that are being shared are the nodes, ports, links and bandwidth of the segment.

The shared resources, for example bandwidth capacity, should be reserved in partitions according to the different SMPGs at the particular segment.

```
   A------B-------C     D------E
    \            /     /        \
     \          /     /          \
      F---G----H-----J------K-----L
         /          /        \     \
        /          M----------N     \
       /                             \
      V-------W-------X-------Y-------Z
```

Figure 2: Shared Mesh Protection Groups

To further clarify, consider the mesh network in Figure 2. In this figure we have the following working paths and corresponding protection paths:

Wx	working path	protection path
W1	A-B-C	A-F-G-H-C
W2	D-E	D-J-K-L-E
W3	M-N	M-J-K-N
W4	V-W-X-Y-Z	V-G-H-J-K-L-Z

In this network we would define three SMPG - characterized by the three shared segments -

o S1 segment G-H - shared by W1 and W4

o S2 segment J-K - shared by W2, W3, and W4

o S3 segment K-L - shared by W2 and W4

The shared segment is always the smallest segment that is shared by multiple protection paths. Therefore, even though segment J-K-L is shared by W2 and W4, we split this into two shared segments - J-K and K-L, since W3 also shares the resources of segment J-K.

In addition, this demonstrates that a single working path may be a member of a number of SMPGs. Also a single SMPG may include more than two working paths.

3.2. Shared Start and End Nodes

For the sake of the discussion of the SMP operation, we designate the two end nodes of the shared protection segment as a Shared Start Node (SSN) and Shared End Node (SEN). To simplify the discussion, this designation is based on referencing the protection path as a pair of unidirectional LSPs.

A SSN is the first node of a unidirectional shared protection segment. For example, in Figure 1, node P is a SSN on unidirectional protection paths A-P-Q-R-E and V-P-Q-R-Z. SSN may act as a Maintenance Entity Group Intermediate Point (MIP) for each protection path sharing the same protection resources.

Similarly, a SEN is defined as the last node of a unidirectional shared protection segment (for example, node R on unidirectional protection paths A-P-Q-R-E and V-P-Q-R-Z in Figure 1). A SEN acts as a MIP on each protection path that shares the protection resource.

Both SEN and SSN are involved in coordinating the use of the unidirectional shared protection segment during the shared mesh

protection operation.

Table 1 summarizes the relationship between SSN and SEN of the shared protection segment and protection paths sharing it as illustrated in Figure 1.

Table 1: SSN/SEN in Figure 1

Protection paths	Shared protection segment	SSN	SEN
A-P-Q-R-E, V-P-Q-R-Z	P-Q-R	P	R
E-R-Q-P-A, Z-R-Q-P-V	R-Q-P	R	P

Figure 3 shows a more complex example of the shared mesh protection domain. Three working paths ABC, DEF, and GHJ are protected by the protection paths APQC, DRSF, and GPQRSJ, respectively. Table 2 summarizes the relationship between SSN and SEN of the shared protection segments and the related protection paths in the protection domain illustrated in Figure 3.

```
A------B------C  D------E------F
 \           /    \           /
  \         /      \         /
   \       /        \       /
    P-----Q----------R-----S
   /                        \
  /                          \
 /                            \
G--------------H---------------J
```

Figure 3: A More Complex Mesh Protection Example

Table 2: SSN/SEN in Figure 3

Protection paths	Shared protection segment	SSN	SEN
A-P-Q-C, G-P-Q-R-S-J	P-Q	P	Q
C-Q-P-A, J-S-R-Q-P-G	Q-P	Q	P
D-R-S-F, G-P-Q-R-S-J	R-S	R	S
F-S-R-D, J-S-R-Q-P-G	S-R	S	R

3.3. SMP protocol communication channel

The MPLS-TP Framework [RFC5921] defines the concept of a Sub-Path Maintenance Entity (SPME) and together with [RFC5586] defines the use of the Generic Associated Channel (G-ACh) for communication of MPLS-TP control protocols between the endpoints of a maintenance entity. While the usual utility of a SPME is to allow tunneling of transport traffic while monitoring the segment with in-band connectivity verification messages, it is possible to use concept of a SPME to describe a LSP that is dedicated to carry a control protocol over the G-ACh between the endpoints of the shared protection segment and the endpoints of the protection paths within the SMPG.

For example, referring to the network in Figure 3, we would configure the following SPME (without identifying the intermediate nodes):

A-P, G-P, P-Q, Q-C and Q-J (for SMPG 1 sharing P-Q), and
D-R, G-R, R-S, S-F and S-J (for SMPG 2 sharing R-S).

These SPMEs are bidirectional LSPs that are not used to carry any data traffic, but only the SMP protocol traffic described in Section 4.

The communication channel between the SSN and SEN of the shared protection segment and between themselves and the endpoints of the protection paths within the SMPG is to coordinate the allocation of the shared segment to a single protection path during a protection switching condition. This process is described more fully in Section 3.6

3.4. Network planning for SMP

Shared mesh protection will typically be dependent upon careful network planning. This includes:

- Preparing the working and protection paths for the different services that require protection.

- Determining which working paths are disjoint and so will not be subject to common failures. It should be clear that working paths within the same SRLG should not be included in the same SMPG.

- Identifying which protection paths share network resources and can constitute a SMPG. Signaling or configuring the proper path information for the shared segment endpoints to allow for communication between the corresponding endpoints of the shared segment and the protection path.

- o Assigning Protection Switching Priority and a path identifier for each working path within a shared mesh protection domain.

- o Ensuring that working paths of high Protection Switching Priority do not share resources on their protection paths in such a way that would mean that one of them could be unprotected.

- o Enabling the necessary shared mesh protection functions at the endpoints of the shared protection segments. This includes preparing the different SPME used for communication between the corresponding endpoints of the shared protection segments and the protection paths, as well as between the endpoints of the shared protection segment.

Note that some control plane features of GMPLS may be used to dynamically configure shared mesh protection. These features are out of scope for this document which focuses on the operation of shared mesh protection switching once it has been configured.

3.5. Preemption and race conditions

In the normal operation of SMP, when a working path triggers a protection switch, and requests allocation of the shared resources, the SMP process should verify that the resources are available and allocate them to the requesting protection path. There are some cases where the determination of the availability is not simply determined.

Within the SMP domain, there is a need to define a "Protection Switching Priority" for each working path. This Protection Switching Priority will be used to determine the use of the shared protection resources in cases of possible preemption. When the shared resources are in use protecting the traffic of a failed working path and a second working path fails, the SMP process should compare the Protection Switching Priority of the two working paths and if the priority of the second path is higher than the priority of the currently protected traffic, then this second path will preempt the currently protected traffic. If the second path has a lower or equal priority to the currently protected traffic, then the second path is locked-out of the protection resources.

The Protection Switching Priority may be provisioned by the network management system or configured by some other mechanism that is outside the scope of this document.

There is an additional case where the SMP process needs to make a determination of which working path should be allowed to be protected using the shared resources. This is the case of multiple working

paths triggering a protection switch virtually simultaneously. This may result in a race condition where the two endpoints of the shared protection segment ostensibly receive requests from two different working paths. By default, working paths with equal priority result in first-come first-served recovery. If multiple working paths request protection switching simultaneously, a pre-defined identifier assigned to each working path in the SMP domain MUST be used to determine the priority among them. The definition of the identifier is for further study.

3.6. SMP Overview

When a protection switching trigger is activated on any of the working paths within a SMPG, then the local linear protection mechanism (in 1:1 protection mode) should cause a protection switch. If, as a result of the protection switch action, there is a need to transmit working data on the protection path then the endpoint of LP domain should inform the endpoint of the shared protection segment of the allocation of the shared resources.

At this point, the shared segment endpoints should notify all of the other protection paths in the SMPG that the resources have been allocated, which could affect the linear protection actions relative to future triggers.

3.6.1. LP Protocol extensions for shared protection

The shared mesh protection mechanism is designed to fully utilize the existing end-to-end LP switching on the working paths. These LP domains SHALL operate in revertive mode. The LP protocol should use the normal procedures for LP without any changes except support for the following additional functionalities:

- o Function to generate a protection switching event message to the SEN when a switching trigger occurs at the end-to-end LP domain. Switching to the protection path or reverting to the working path should be notified.

- o Function to take a Lockout of Protection (LoP) request message from the SEN, and incorporate it as the Lockout of Protection (LoP) command assertion or clearance.

- o Function to acknowledge the SEN when the LP domain completes the LoP operation.

3.6.2. Notifying protection switching event

If the endpoint of a working path detects a switching trigger, it triggers the protection switching and exchanges LP switching protocol messages with far endpoint. This operation is independent of the SMP switching mechanism specified in this document.

At the same time, for the operation of SMP, the protection path endpoint notifies its protection switching event to SENs by sending a "Protection Switching (PS) Event" message.

The PS Event message MUST be transmitted immediately when an endpoint of the end-to-end LP domain changes its selector position either from working to protection or vice versa. The event message SHALL be transmitted over the SPME that is configured between the protection path endpoint and the SEN, using the G-ACh. When bidirectional protection switching is being used by the working path, both endpoints will transmit the event messages to their corresponding SENs using the properly configured SPME. When unidirectional protection switching is being used and a unidirectional failure is detected, only the detecting endpoint will send the messages to its corresponding SENs.

The endpoint of the protection path that is becoming active (or released) sends the messages directly to each SEN. This requires that N messages are sent, where N is the number of SMPG that the working path is a member of. This, of course, implies that the endpoints are pre-configured with knowledge of all SENs associated with the SMPG.

3.6.3. Requesting lockout of protection

When a SEN receives the PS Event message notifying that protection switching to the protection path has begun in an end-to-end LP domain and that the shared resources are to be allocated, it compares the Protection Switching Priority of the working path notifying the event with those of other LP domains in the same SMPG.

The SEN determines which of the LP domains (within the SMPG) have a lower or equal priority to that of the notifying LP domain. The SEN then sends a "Lockout of Protection (LoP) Request" message to the endpoints of these protection paths that is equivalent to a "Lockout of Protection" operator command. This prevents any protection switching actions in those LP domains. For those LP domains having higher priorities no request is transmitted and those LP domains may continue to perform protection switching actions which they require.

When a protection path endpoint receives the LoP Request message from

an SEN, it SHOULD react as if a LoP command was received, according to the actions dictated by the LP protocol. Since the LoP command has the highest priority in the LP switching protocol, it will inhibit any further protection switching in the LP domain.

If the LP domain that received the LoP Request message is currently transmitting traffic on the protection path, it SHALL immediately stop transmitting the traffic on the protection path and release the allocated resources.

When the protection path endpoint completes LoP operation, it SHOULD immediately reply with a "LoP Acknowledgement" message to inform the completion of the LoP operation to the SEN.

To minimize potential congestion that may occur when a protection path having higher priority pre-empts other protection paths having lower priorities, the SEN SHOULD block forwarding traffic from the protection paths having lower priorities until it receives the LoP Ack message from the endpoints of those protection paths.

When a SEN receives a PS Event message indicating that the shared protection resources are being released, i.e. the LP domain is reverting to normal state, it sends a LoP Request message to the endpoints of all the protection paths in the SMPG that were previously locked (i.e. those with equal or lower priority) to clear the LoP command. The endpoint of the protection path that receives this message SHALL react as if a Clear command was received.

3.6.4. Resolving race conditions

As was pointed out in Section 3.5 there are some cases, in particular in unidirectional protection switching triggers, of simultaneous protection switching that could cause race conditions. In these use-cases there is a need for the two end nodes of the shared protection segment, i.e. the SEN and the SSN, to coordinate the selection of the LP domain that will be allocated the shared protection resources.

For this purpose, additional messages are defined that are transmitted on the SPME that is defined between the end nodes of the shared protection segment. When a SEN receives a PS Event message from a LP domain indicating that protection switching to the protection path has begun, it SHALL send a "Resource Allocation (RA) Notification" message to the SSN that the resources have been allocated, with an indication of the working path identifier. This allocation needs to be confirmed for cases where both end nodes report allocation to different working path identifiers.

The race condition can occur only when more than one protection paths

are configured to have the same Protection Switching Priority within a SMP domain.

When the SSN receives the RA Notification message from the SEN, it first checks whether it has received a PS Event message from an endpoint of the other LP domain having the same Protection Switching Priority that corresponds to the LP domain sending the RA notification message.

If both have the same priority, the SSN compares the working path identifier and sends an RA Ack message to the SEN only when it is determined that the working path identifier contained in the RA Notification message to have been allocated the shared protection resources. Each working path or LP domain has a unique identifier within a SMP domain and rules for deciding the priority will be defined in later.

The SEN does not perform the LoP request procedure described in Section 3.6.3 until it receives an RA Ack from the SSN. This results in the overall protection switching time to be increased. To avoid this, it is RECOMMENDED to configure none of the working paths sharing the protection segment in a SMP domain to have the same Protection Switching Priority.

4. Protocol

4.1. PDU Format

The shared mesh protection protocol messages MUST be sent over a G-ACh as defined in [RFC5586].

The shared mesh protection protocol messages are as follows:

- o Protection Switching (PS) Event message [sent from protection path endpoint to SEN]

- o Lockout of Protection (LoP) Request message [sent from SEN to protection path endpoint]

- o Lockout of Protection (LoP) Acknowledgement message [sent from protection path endpoint to SEN]

- o Resource Allocation (RA) Notification message [sent from SEN to SSN]

- o Resource Allocation (RA) Acknowledgement message [sent from SSN to SEN]

The channel type in ACH is used to indicate that the message is a SMP protocol message. The protocol message MUST follow the ACH.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 1|Version|   Reserved    | Channel Type = Shared Mesh P. |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Shared Mesh Protection Protocol Message             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Shared mesh protection protocol message header

The SMP protocol message format is defined as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Version|  Mode |  Type | ST|     Reserved      |       ID      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: Shared mesh protection protocol message format

Each protocol message includes the following fields:

o Version - Version indicator
    0x0   - reserved
    0x1   - This version
    0x2~0xF - reserved

o Mode - SMP switching and operation mode
    0x0   - Bidirectional
    0x1   - Unidirectional
    0x2~0xF - reserved

o Type - SMP protocol message type indicator
    0x0   - reserved
    0x1   - PS Event
    0x2   - LoP Request
    0x3   - LoP Acknowledgement
    0x4   - RA Notification
    0x5   - RA Acknowledgement
    0x6~0xF - reserved

o ST - SMP protocol message sub-type indicator

```
0x0     - Switch to working (for PS Event),
          Clear LoP (for LoP Request)
0x1     - Switch to protection (for PS Event),
          Assert LoP (for LoP Request)
0x2~0x3 - reserved
```

o ID - This is either the identifier of the LP domain that is sending the message or the working path that was allocated to the resources (dependent upon the message).

```
0x0      - reserved
0x1~0xFE - ID value
0xFF     - reserved
```

o Reserved - reserved fields

All reserved bits SHOULD be zero upon transmission, and MUST be ignored on reception.

4.2. Message Transmission

A new message must be transmitted immediately. The first three messages should be transmitted as fast as possible so that fast protection switching is possible even if one or two messages are lost or corrupted. The interval of the first three messages should be less than 3.3ms. Messages after the first three should be transmitted with the interval of 5 seconds.

If no valid message is received, the last valid received information remains applicable.

5. Operation of Shared Mesh Protection

This section illustrates the operation of the SMP protocol based on the example illustrated in Figure 3 and the following assumptions:

o The SMP domain consists of the following end-to-end LP domains (LPDs):

  * LPD1: Working path ABC (W1) / Protection path APQC (P1)

  * LPD2: Working path GHJ (W2) / Protection path GPQRSJ (P2)

  * LPD3: Working path DEF (W3) / Protection path DRSF (P3)

o The SMP domain includes the following SMPG:

  * S1: LPD1 & LPD2 (Shared protection segment - PQ)

  * S2: LPD3 & LPD2 (Shared protection segment - RS)

o Protection Switching Priority is LPD1 > LPD2 > LPD3 (i.e. LPD1 has the highest priority.)

o All working paths are protected by 1:1 bidirectional protection switching.

If a unidirectional failure occurs on W2 in the direction from node H to node G as shown in Figure 6, SMP will perform the following:

a. Node G detects the failure, and initiates linear protection switching for the failed W2.

b. At the same time, node G transmits the PS Event message notifying the SENs of the shared protection segments for S1 & S2, i.e. P and R, that a protection switching event occurred to node G.

c. SEN P compares the Protection Switching Priority of LPD2 with those of other members of S1, i.e. LPD1. In this example, since the priority of LPD1 is higher than LPD2, SEN P does not send any message to node A.

d. SEN R compares the Protection Switching Priority of LPD2 with those of other members of S2, i.e. LPD3. In this example, as the priority of LPD3 is lower than LPD2, SEN R sends the RA Notification message to the SSN S, blocks forwarding of P3 and sends the LoP Request message requesting the assertion of LoP to node D.

e. SSN S does not process the RA Notification message. (Since in this example, all the LP domains are configured to have different Protection Switching Priorities.)

f. Node D takes the LoP Request message as input to the LP switching, and follows the LP procedure to process the end-to-end LoP command. After completion of the LoP operation, node D sends the LoP Ack message to SEN R.

g. SEN R unblocks forwarding of P3 upon receiving the LoP Ack message from node D.

h. Since LPD2 operates in 1:1 bidirectional protection switching mode, node J performs the switching operations (i.e. switches its

bridge and selector state) to synchronize with node G, and also transmits the PS Event message to node S and Q, which are SENs for G->H->J. Using a parallel procedure to that described in steps c & d, SEN S sends the LoP Request message to node F while the SEN Q does not take an action to node C.

```
                   W1                     W3
          ==A======B======C==    ==D======E======F==
             \           /          \           /
              \   LPD1  /            \   LPD3  /
               \       /              \       /      == : Normal traffic
                P=====Q================R=====S
               //                            \\
              //             LPD2             \\
             //                                \\
          ==G------xx---------H------------------J==
               SF on G<-H     W2
```

Figure 6: Shared Mesh Protection Example 1

Figure 7 shows a progression from Figure 6. While LPD2 is in protecting state with its traffic transported on protection path P2, another unidirectional failure occurs on W1 in the direction from node B to node A.

In this case, the shared mesh protection will operate as follows:

a. Node A detects the failure, and initiates the linear protection switching for the failed W1.

b. At the same time, node A transmits the PS Event message notifying SEN for S1, i.e. node P, that a protection switching event occurred.

c. SEN P compares the Protection Switching Priority of LPD1 with those of the other members in S1, in this case LPD2. In this example, since the priority of LPD2 is lower than LPD1, SEN P sends the RA Notification message requesting the assertion of LoP to node G.

d. SSN Q does not process the RA Notification message. (Since in this example, all the LPDs are configured to have different Protection Switching Priorities.)

e. Node G accepts the LoP Request message as input to linear protection switching, and follows LP procedure to process the LoP command. When LPD2 is forced to lockout its protection path P2,

it may try to find another available path. m:n protection or other recovery mechanism may be used for this, but this discussion is out of scope for this document. After completion of the LoP operation, node G sends the LoP Ack message to SEN P.

f. SEN P unblocks forwarding of P2 upon receiving the LoP Ack message from node G.

g. As node G changes its bridge and selector states from protection to working, it will transmit the PS Event message to the SENs of S1 & S2, i.e. P & R, notifying that the shared protection resources should be released.

h. SEN P compares the Protection Switching Priority of LPD2 with the other members of S1, i.e. LPD1, and does not transmit any message to node A, but SEN R sends the LoP Request message to request the clear of LoP to node D, after comparing the Protection Switching Priorities of the members of S2.

i. Node D accepts the message as input to the linear protection switching, and follows the LP procedures to clear the LoP command.

```
              SF on
               A<-B W1                    W3
           ==A-xx---B------C==   ==D======E======F==
             \\           //        \           /
              \\   LPD1  //          \   LPD3  /
               \\       //            \       /      == : Normal traffic
                 P=====Q---------------R-----S
                /                             \
               /              LPD2             \
              /                                 \
           ==G------xx---------H----------------J==
                SF on G<-H     W2
```

Figure 7: Shared Mesh Protection Example 2

NOTE: Examples for race condition to be provided in the next version.

6. Manageability Considerations

To be added in future version.

7. IANA Considerations

To be added in future version.

8. Security Considerations

To be added in future version.

9. References

9.1. Normative References

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

[RFC5654] Niven-Jenkins, B., Nadeau, T., and C. Pignataro, "Requirements for the Transport Profile of MPLS", RFC 5654, April 2009.

9.2. Informative References

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

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

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

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

[RFC5085] Nadeau, T. and C. Pignataro, "Pseudo Wire (PW) Virtual Circuit Connectivity Verification ((VCCV): A Control Channel for Pseudowires", RFC 5085, Dec 2007.

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

[RFC4428] Papadimitriou, D. and E. Mannie, "Analysis of Generalized Multi-Protocol Label Switching (GMPLS) based Recovery Mechanisms (including Protection and Restoration) Recovery (Protection and Restoration)", RFC 4428, March 2006.

[G.808.1] SG15, "Generic Protection Switching - Linear trail and

subnetwork protection", ITU-T G.808.1, Feb 2010.

Authors' Addresses

Tae-sik Cheung
ETRI
161 Gajeong
Yuseong, Daejeon 305-700
South Korea

Phone: +82 42 860 5646
Email: cts@etri.re.kr

Jeong-dong Ryoo
ETRI
161 Gajeong
Yuseong, Daejeon 305-700
South Korea

Phone: +82 42 860 5384
Email: ryoo@etri.re.kr

Yaacov Weingarten
Nokia Siemens Networks
3 Hanagar St. Neve Ne'eman B
Hod Hasharon, 45241
Israel

Phone: +972-54-220 0977
Email: wyaacov@gmail.com

Nurit Sprecher
Nokia Siemens Networks
3 Hanagar St. Neve Ne'eman B
Hod Hasharon, 45241
Israel

Email: nurit.sprecher@nsn.com

Daniel King
Old Dog Consulting
United Kingdom

Email: daniel@olddog.co.uk