Internet Engineering Task Force   Yimin Shen
Internet-Draft   Juniper Networks
Intended status: Standards Track   Rahul Aggarwal
Expires: July 30, 2016   Arktan, Inc
                                                    Wim Henderickx
                                                    Alcatel-Lucent
                                                    Yuanlong Jiang
                                                Huawei Technologies
                                                   January 27, 2016

                    PW Endpoint Fast Failure Protection
                 draft-ietf-pals-endpoint-fast-protection-02

Abstract

   This document specifies a fast mechanism for protecting pseudowires
   against egress endpoint failures, including egress attachment circuit
   failure, egress PE failure, multi-segment PW terminating PE failure,
   and multi-segment PW switching PE failure. Operating on the basis of
   multi-homed CE, redundant PWs, upstream label assignment and context
   specific label switching, the mechanism enables local repair to be
   performed by the router upstream adjacent to a failure. The router
   can restore a PW in the order of tens of milliseconds, by rerouting
   traffic around the failure to a protector through a pre-established
   bypass tunnel. Therefore, the mechanism can be used to reduce
   traffic loss before global repair reacts to the failure and the
   network converges on the topology changes due to the failure.

Status of This Memo

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

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF). Note that other groups may also distribute
   working documents as Internet-Drafts. The list of current Internet-
   Drafts is at http://datatracker.ietf.org/drafts/current/.

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

   This Internet-Draft will expire on July 30, 2016.

Copyright Notice

Table of Contents

1. Introduction

   Per RFC3985, RFC4447 and RFC5659, a pseudowire (PW) or PW segment can
   be thought of as a connection between a pair of forwarders hosted by
   two PEs, carrying an emulated layer-2 service over a packet switched
   network (PSN). In the single-segment PW (SS-PW) case, a forwarder
   binds a PW to an attachment circuit (AC). In the multi-segment PW
   (MS-PW) case, a forwarder on a terminating PE (T-PE) binds a PW
   segment to an AC, while a forwarder on a switching PE (S-PE) binds
   one PW segment to another PW segment. In each direction between the
   PEs, PW packets are transported by a PSN tunnel, which is also called
   a transport tunnel.

   In order to protect the layer-2 service against network failures, it
   is necessary to protect every link and node along the entire data
   path. For the traffic in a given direction, this include ingress AC,
   ingress (T-)PE, intermediate routers of transport tunnel, S-PEs,
   egress (T-)PE, and egress AC. To minimize service disruption upon a
   failure, it is also desirable that each of these components is
   protected by a fast protection mechanism based on local repair. Such
   mechanism generally involves a bypass path that is pre-computed and
   pre-installed in the data plane on the router upstream adjacent to an
   anticipated failure. The bypass path has the property that it can
   guide traffic around the failure, while remaining unaffected by the
   topology changes resulting from the failure. When the failure
   occurs, the router can invoke the bypass path to achieve fast
   restoration for the service.

   Today, fast protection against ingress AC failure and ingress (T-)PE
   failure can be achievable by using a multi-homed CE and redundant
   ACs, such as multi-chassis link aggregation group (MC-LAG). Fast
   protection against failure of intermediate router of transport tunnel
   can be achievable through RSVP fast-reroute [RFC4090] or IP/LDP fast-
   reroute [RFC5714, RFC5286]. However, there is a lack of equivalent
   mechanism against egress AC failure, egress (T-)PE failure, and S-PE
   failure. For these failures, service restoration has to rely on
   global repair or control plane repair. Global repair normally
   involves ingress CE or ingress (T-)PE switching traffic to another
   fully functional path, based on remote failure detection via PW
   status notification, end-to-end OAM, etc. Control plane repair
   relies on control protocols to converge on the topology changes due
   to a failure. Compared to local repair, these mechanisms are
   relatively slow in reacting to a failure and restoring traffic.

This document is intended to serve the above need. It specifies a
fast protection mechanism based on local repair to protect PWs
against the following endpoint failures.

a. Egress AC failure.

b. Egress PE failure: Link or node failure of an egress PE of an SS-
    PW, or a T-PE of an MS-PW.

c. Switching PE (S-PE) failure: Link or node failure of an S-PE of
    an MS-PW.

The mechanism is applicable to LDP signaled PWs. It is relevant to
networks with redundant PWs and multi-homed CEs. It is designed on
the basis of MPLS upstream label assignment and context-specific
label switching [RFC5331]. Fast protection refers to its ability to
restore traffic in the order of tens of milliseconds. Compared with
global repair and control plane repair, this mechanism can provide
faster service restoration. However, it is intended to complement
those mechanisms, rather than replacing them in any fashion.

2. Specification of Requirements

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.

3. Reference Models for Egress Endpoint Failures

This document refers to the following topologies to describe egress
endpoint failures and protection procedures.

3.1. Single-Segment PW

```
                 |<-------------- PW1 --------------->|

               - PE1 ------------- P1 --------------- PE2 -
              /                                            \
             /                                              \
          CE1                                                CE2
             \                                              /
              \                                            /
               - PE3 ------------- P2 --------------- PE4 -

               |<-------------- PW2 --------------->|
```
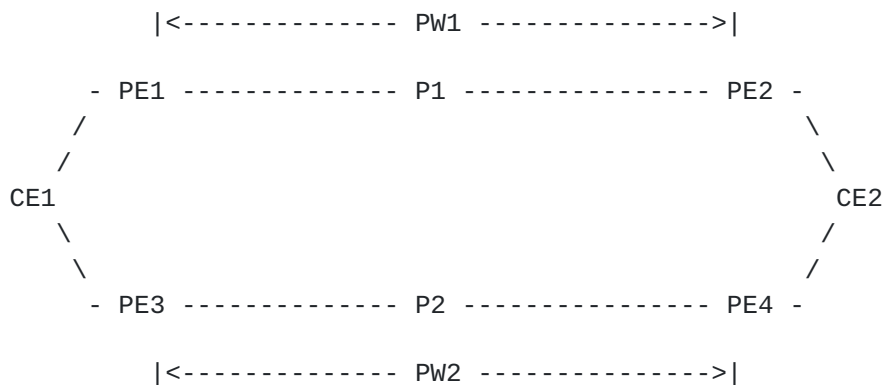
Figure 1

In Figure 1, the IP/MPLS network consists of PE and P routers. It
provides an emulation of a layer-2 service between CE1 and CE2. Each
CE is multi-homed via two ACs to two PEs. This forms two divergent
paths between the CEs. The first path uses PW1 between PE1 and PE2,
and the second path uses PW2 between PE3 and PE4. The transport
tunnels of the PWs and other links between the routers are not shown
in this figure for clarity.

In general, a CE may operate the ACs in two modes when sending
traffic to the remote CE, i.e. active-standby mode and active-active
mode.

o   In the active-standby mode, the CE chooses one AC as active AC and
    the corresponding path as active path, and uses the other AC as
    standby AC and the corresponding path as standby path. The CE
    only sends traffic on the active AC as long as the active path is
    operational. The CE will only send traffic on the standby AC
    after it detects a failure of the active path. Note that the CE
    may receive traffic on the active or standby AC, depending on
    whether the remote CE chooses the same active path for the traffic
    of the reverse direction. In this document, even if both CEs
    choose the same active path, each CE should still anticipate
    receiving traffic on a standby AC, because the traffic may be
    redirected to the standby path by the fast protection mechanism.

o   In the active-active mode, the CE treats both ACs and their
    corresponding paths as active, and sends traffic on both ACs in a
    load balance fashion. In the reverse direction, the CE may
    receive traffic on both ACs.

For either mode, when considering the traffic flowing in a given
direction over an active path, this document views the ACs, PEs and
PWs to serve primary or backup roles. In particular, the ACs, PEs
and PW along this active path are primary, while those along the
other path are backup. Note that in the active-active mode, the
backup path is an active path by itself, carrying its own share of
traffic while protecting the other active path.

For Figure 1, the following roles are assumed for the traffic going
from CE1 to CE2 via PW1.

    Primary ingress AC: CE1-PE1

    Primary ingress PE: PE1

    Primary PW: PW1

    Primary egress PE: PE2

Primary egress AC: PE2-CE2

Backup ingress AC: CE1-PE3

Backup ingress PE: PE3

Backup PW: PW2

Backup egress PE: PE4

Backup egress AC: PE4-CE2

Based on this schema, this document describes egress endpoint failures and the fast protection mechanism on the per-active-path and per-direction basis. In this case, an egress AC failure refers to the failure of the AC PE2-CE2, and an egress node failure refers to the failure of PE2. The ultimate goal is that when a failure occurs, the traffic should be locally repaired, so that it can eventually reach CE2 via the backup egress PE (PE4) and the backup egress AC (PE4-CE2).

Subsequent to the local repair, either the current active path should heal after control plane converges on the new topology, or the ingress CE should switch traffic from the primary path to the backup path, depending on the failure scenario. In the latter case, the ingress CE may perform the path switchover triggered by end-to-end OAM (in-band or out-band), PW status notification, CE-PE control protocols (e.g. LACP), etc. In the active-standby mode, this will promote the standby path to new active path. In the active-active mode, it will make the other active path carry all the traffic between the two CEs. In any case, this phase of restoration falls into the control plane repair and global repair category, and hence is out of the scope of this document. The purpose of the fast protection mechanism in this document is to reduce traffic loss before this phase of restoration takes place.

Note that an egress endpoint failure of the traffic of a given direction may be detected by the egress CE as an ingress endpoint failure for the traffic of the reverse direction, except when the failure is on a link of the primary egress PE within the PSN, or when the traffic of the reverse direction takes a different active path. If the CE can detect the failure, it may protect the traffic of the reverse direction by switching it to the backup path. However, this is categorized as ingress endpoint failure protection, and hence is not handled by this mechanism.

Figure 2 shows another possible scenario, where CE1 is single-homed to PE1, while CE2 remains multi-homed to PE2 and PE4. From the

perspective of egress endpoint protection for the traffic going from
CE1 to CE2 over PW1, this scenario is not much different than
Figure 1.

```
                 |<-------------- PW1 --------------->|

                  ------------ P1 ---------------- PE2 -
                 /                                       \
                /                                         \
          CE1 -- PE1                                       CE2
                \                                         /
                 \                                       /
                  ------------ P2 ---------------- PE4 -

                 |<------------- PW2 --------------->|
```
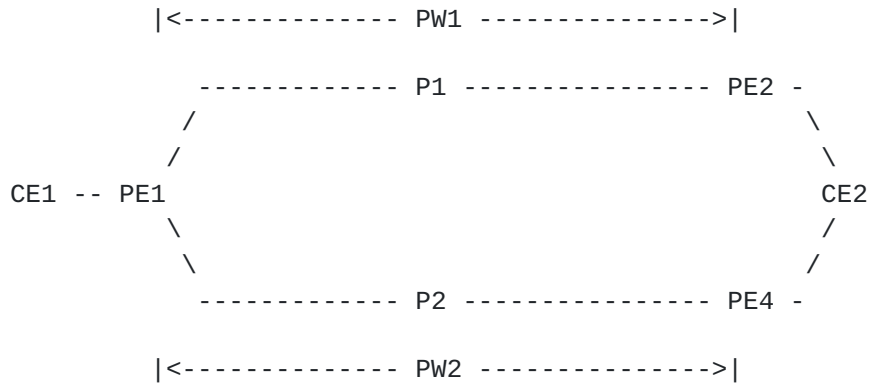
                             Figure 2

For clarity, primary egress AC, primary egress PE, backup egress AC,
and backup egress PE may simply be referred to as primary AC, primary
PE, backup AC, and backup PE, respectively, when the context of a
discussion is egress endpoint.

3.2. Multi-Segment PW

```
                 |<-------------- PW1 --------------->|
                 |<----- SEG1 ----->|<----- SEG2 ----->|

               - TPE1 -------------- SPE1 -------------- TPE2 -
              /                                               \
             /                                                 \
          CE1                                                   CE2
             \                                                 /
              \                                               /
               - TPE3 -------------- SPE2 -------------- TPE4 -

                 |<----- SEG3 ----->|<----- SEG4 ----->|
                 |<-------------- PW2 --------------->|
```
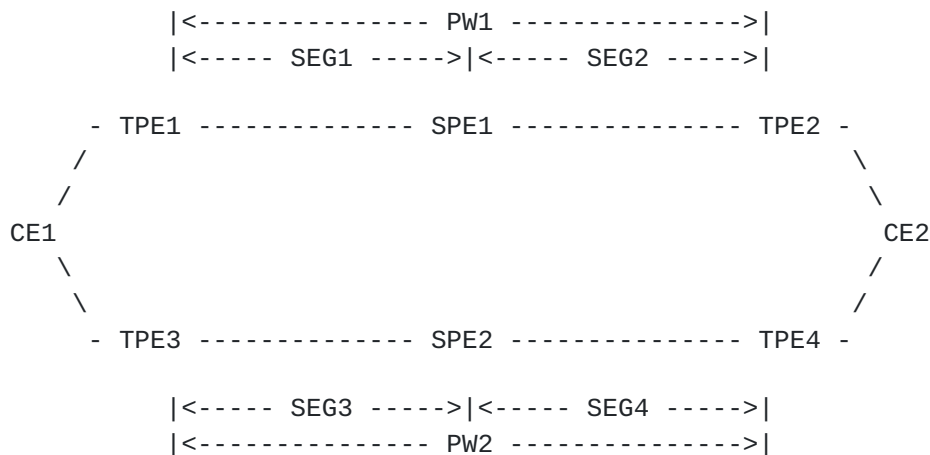
                             Figure 3

Figure 3 shows a topology that is similar to Figure 1 but in an MS-PW
environment. PW1 and PW2 are both MS-PWs. PW1 is established
between TPE1 and TPE2, and switched between segments SEG1 and SEG2 at
SPE1. PW2 is established between TPE3 and TPE4, and switched between
segments SEG3 and SEG4 at SPE2. CE1 is multi-homed to TPE1 and TPE3.
CE2 is multi-homed to TPE2 and TPE4. The transport tunnels of the PW
segments are not shown in this figure for clarity.

In this document, the following primary and backup roles are assigned for the traffic going from CE1 to CE2:

   Primary ingress AC: CE1-TPE1

   Primary ingress T-PE: TPE1

   Primary PW: PW1

   Primary S-PE: SPE1

   Primary egress T-PE: TPE2

   Primary egress AC: TPE2-CE2

   Backup ingress AC: CE1-TPE3

   Backup ingress T-PE: TPE3

   Backup PW: PW2

   Backup S-PE: SPE2

   Backup egress T-PE: TPE4

   Backup egress AC: TPE4-CE2

In this case, an egress AC failure refers to the failure of the AC TPE2-CE2. An egress node failure refers to the failure of TPE2. An S-PE failure refers to the failure of SPE1.

For consistency with the SS-PW scenario, primary T-PEs and a primary S-PEs may simply be referred to as primary PEs in this document, where specifics are not required. Similarly, backup T-PEs and backup S-PEs may be referred to as backup PEs.

4. Theory of Operation

The fast protection mechanism in this document provides three types of protection for PWs, corresponding to the three types of failures described in Section 1.

a. Egress AC protection

b. Egress (T-)PE node protection

c. S-PE node protection

4.1. Applicability

   The mechanism is applicable to LDP signaled PWs in an environment
   where an egress CE is multi-homed to a primary PE and a backup PE and
   there exists a backup PW, as described in Section 3.   The procedure
   for S-PE node protection is applicable when there exists a backup
   S-PE on the backup PW.

   The mechanism assumes IP/MPLS transport tunnels. In a network where
   transport tunnels may provide ECMP to primary PEs, care should be
   taken to prevent misordered packet delivery during local repair.
   Imagine a scenario where the transport tunnel of a PW traverses a
   router with ECMP to a primary PE, and the ECMP include a direct link
   to the primary PE. Normally the router will attempt to forward PW
   packets in a load balance fashion over the ECMP, including this link.
   In this document, when the link fails, the router will treat the
   event as an egress PE failure, and reroute the portion of traffic on
   the link towards a backup PE. Meanwhile, the rest of the traffic
   will remain on the other ECMP branches to the primary PE. This will
   create a situation where the egress CE receives traffic from both the
   primary PE and the backup PE, which is undesirable if the PW or flows
   within the PW are sensitive to packet misordering. Therefore, the
   mechanism assumes that Control Word (CW) SHOULD be used for PWs and
   flow labels [RFC6391] SHOULD be used for flows within a PW, whenever
   applicable.

   It is also assumed that the mechanism SHOULD be used in conjunction
   with global repair and control plane repair, in such a manner that
   the mechanism temporarily repairs a failed path by using a bypass
   tunnel, and global repair and control plane repair eventually move
   traffic to a fully functional path.

4.2. Local Repair and Protector

   The fast protection ability of the mechanism comes from local repair
   performed by routers upstream adjacent to failures. Each of these
   routers is referred to as a "point of local repair" (PLR). A PLR
   MUST be able to detect a failure by using a rapid mechanism, such as
   physical layer failure detection, Bidirectional Failure Detection
   (BFD) [RFC5880], etc. In anticipation of the failure, the PLR MUST
   also pre-establish a bypass tunnel to a "protector", and pre-install
   a bypass route in the data plane. The bypass tunnel MUST have the
   property that it will not be affected by topology changes due to the
   failure. Upon detecting the failure, the PLR invokes the bypass
   route in the data plane, and reroutes PW traffic to the protector
   through the bypass tunnel. The protector in turn sends the traffic
   to the target CE. This procedure is referred to as local repair.

Different routers may serve as PLR and protector in different
scenarios.

o In egress AC protection, the PLR is the primary PE, and the
   protector is the backup PE (Figure 4).

```
             |<------------- PW1 --------------->|

          - PE1 ------------- P1 --------------- PE2 -
         /                                       PLR  \
        /                                        |     \
    CE1                                     bypass|      CE2
        \                                        |     /
         \                                       |    /
          - PE3 ------------- P2 --------------- PE4 -
                                                 protector

             |<------------- PW2 --------------->|
```

                         Figure 4

o In egress PE node protection, the PLR is the penultimate hop
   router of the transport tunnel of the primary PW, and the
   protector is the backup PE (Figure 5).

```
             |<------------- PW1 --------------->|

          - PE1 ------------- P1 ------- P3 ----- PE2 -
         /                          PLR \         \
        /                                \         \
    CE1                             bypass\         CE2
        \                                  \       /
         \                                  \     /
          - PE3 ------------- P2 --------------- PE4 -
                                                 protector

             |<------------- PW2 --------------->|
```
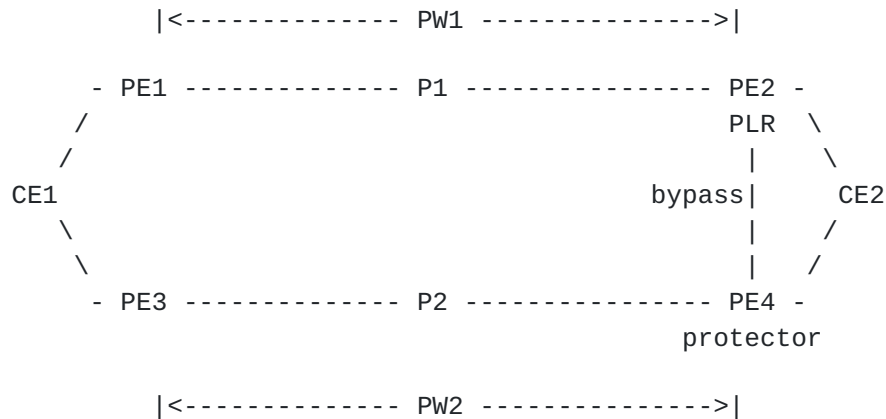
                         Figure 5
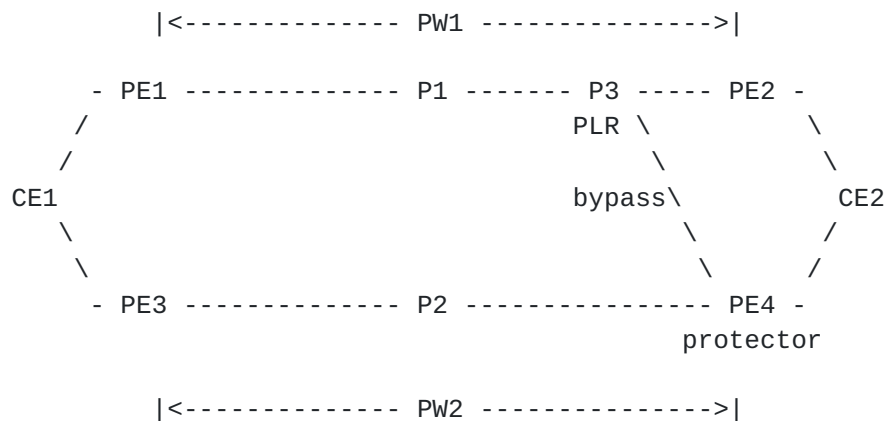
o In S-PE node protection, the PLR is the penultimate hop router of
   the transport tunnel of the primary PW segment, and the protector
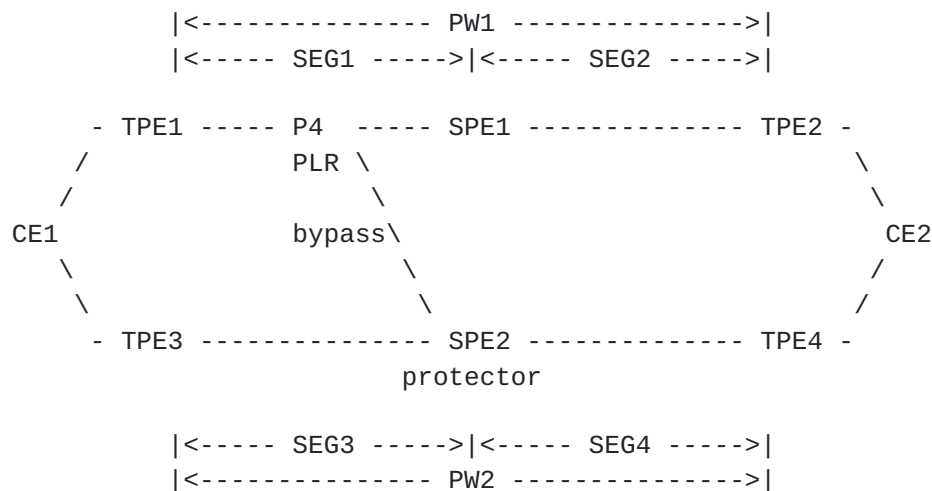   is the backup S-PE (Figure 6).

```
                 |<-------------- PW1 --------------->|
                 |<----- SEG1 ----->|<----- SEG2 ----->|

             - TPE1 ----- P4  ----- SPE1 ------------- TPE2 -
            /              PLR \                              \
           /                    \                              \
        CE1                  bypass\                            CE2
           \                       \                           /
            \                       \                         /
             - TPE3 -------------- SPE2 ------------- TPE4 -
                              protector

                 |<----- SEG3 ----->|<----- SEG4 ----->|
                 |<-------------- PW2 --------------->|
```

                            Figure 6

In egress AC protection, a PLR realizes its role based on
configuration of a "context identifier" introduced in this document
(Section 4.3). The PLR establishes a bypass tunnel to the protector
in the same fashion as a normal PSN tunnel.

In egress PE and S-PE node protection, a PLR is a transit router on
the transport tunnel, and it normally does not have knowledge of the
PW(s) carried by the transport tunnel. In this document, the PLR
simply computes and establishes a node protection bypass tunnel in
the same fashion as the normal IP/MPLS node protection, except that
with the notion of context identifier, the bypass tunnel will be
established from the PLR to the protector (Section 4.6). Conversely,
when the router is no longer a PLR for egress PE or S-PE node
protection due to a change in network topology or the transport
tunnel's path, the router should revert to the role of regular
transit router, including PLR for normal IP/MPLS link or node
protection.

In local repair, a PLR simply switches all the traffic received on
the transport tunnel to the bypass tunnel. This requires that the
protector given by the bypass tunnel MUST be intended for all the PWs
carried by the transport tunnel. This is achieved by the ingress PE
associating a PW with the specific pair of {primary PE, protector}
and mapping the PW to a transport tunnel destined for the same
{primary PE, protector}. The ingress PE MAY map multiple PWs to the
transport tunnel, if they share the {primary PE, protector} in
common.

In local repair, the PLR keeps PW label intact in packets. This
obviates the need for the PLR to maintain bypass routes on a per-PW
basis, and allows bypass tunnel sharing between PWs. On the other

hand, this imposes a requirement on the protector that it MUST be
able to forward the packets based on a PW label that is assigned by
the primary PE, and ensure that the traffic MUST eventually reach the
target CE. From the protector's perspective, this PW label is an
upstream assigned label [RFC5331]. To achieve this, the protector
MUST learn the PW label from the primary PE prior to the failure, and
install proper forwarding state for the PW label in a dedicated label
space associated with the primary PE. During local repair, the
protector MUST perform PW label lookup in this label space.

The previous examples have shown the scenarios where the protectors
are backup (T/S-)PEs. It is also possible that a protector is a
dedicated router to serve such role, separate from the backup (T/
S-)PE. During local repair, the PLR still reroutes traffic to the
protector through a bypass tunnel. The protector then forwards the
traffic to the backup (T/S-)PE, which further forwards the traffic to
the target CE via a backup AC or a backup PW segment. More detail
will be described in Section 4.4.

4.3. Context Identifier

A protector may protect multiple primary PEs. The protector MUST
maintain a separate label space for each primary PE. Likewise, the
PWs terminated on a primary PE may be protected by multiple
protectors, each for a subset of the PWs. In any case, a given PW
MUST be associated with one and only one pair of {primary PE,
protector}.

This document introduces the notion of "context identifier" to
facilitate protection establishment. A context identifier is an
IPv4/v6 address assigned to an ordered pair of {primary PE,
protector}. The address MUST be globally unique, or unique in the
address space of the network where the primary PE and the protector
reside.

4.3.1. Semantics

The semantics of a context identifier is twofold.

o A context identifier identifies a primary PE and an associated
   protector. It represents the primary PE as PW destination on a
   per protector basis. A given primary PE may be protected by
   multiple protectors, each for a subset of the PWs terminated on
   the primary PE. A distinct context identifier MUST be assigned to
   the primary PE and each protector.

   The ingress PE of a PW learns the context identifier of the PW's
   {primary PE, protector} from the primary PE via Interface_ID TLV

[RFC3471, RFC3472] in the LDP Label Mapping message of the PW.
The ingress PE then sets up or resolves a transport tunnel with
the context identifier, rather than a private IP address of the
primary PE, as destination. This destination not only makes the
transport tunnel reach the primary PE, but also conveys the
identity of the protector to the PLR, which MUST use the context
identifier as destination for the bypass tunnel to the protector.
The ingress PE MUST map only the PWs terminated by the exact
primary PE and protected by the exact protector to the transport
tunnel.

o A context identifier indicates the primary PE's label space on the
   protector. The protector may protect PWs for multiple primary
   PEs. For each primary PE, it MUST maintain a separate label space
   to store the PW labels assigned by that primary PE. It associates
   a PW label with a label space via the context identifier of the
   {primary PE, protector}, as below.

   In addition to the normal LDP PW signaling, the primary PE MUST
   have a targeted LDP session with the protector, and advertise PW
   labels to the protector via LDP Label Mapping messages
   (Section 6). The primary PE MUST attach the context identifier to
   each message. Upon receiving the message, the protector MUST
   install the advertised PW label in the label space identified by
   the context identifier.

   When a PLR sets up or resolves a bypass tunnel to the protector,
   it MUST use the context identifier rather than a private IP
   address of the protector as destination. The protector MUST use
   the bypass tunnel, either the MPLS tunnel label or IP tunnel
   destination address, as the pointer to the corresponding label
   space. The protector MUST forwards PW packets received on the
   bypass tunnel based on label lookup in that label space.

4.3.2. IGP Advertisement and Path Computation

   Using a context identifier as destination for both transport tunnel
   and bypass tunnel requires coordination between the primary PE and
   the protector in IGP advertisement of the context identifier in
   routing domain and TE domain. The context identifier should be
   advertised in such a way that all the routers on the tunnels MUST be
   able to independently reach the following common view of paths.

   o The transport tunnel MUST have the primary PE as path endpoint.

   o The bypass tunnel MUST have the protector as path endpoint. In
      egress PE and S-PE node protection, the path MUST avoid the
      primary PE.

There are generally two categories of approaches to achieve the above.

o The first category does not require an ingress PE or a PLR to have knowledge of the PW egress endpoint protection schema. It does not require any IGP extension for context identifier advertisement. A context identifier is advertised by the primary PE and the protector as an address reachable via both routers. The ingress PE and the PLR can compute paths by using a normal method, such as Dijkstra, CSPF (constrained shortest path first), LFA [RFC5286] and MRT [IP-LDP-FRR-MRT]. One example is to advertise a context identifier as a virtual proxy node connected to the primary PE and the protector, with the link between the proxy node and the primary PE having a more preferable IGP and TE metric than the link between the proxy node and the protector. The transport tunnel will follow the shortest path or a TE path to the primary PE, and be terminated by the primary PE. The PLR will no longer view itself as a penultimate hop of the transport tunnel, but rather two hops away from the proxy node, via the primary PE. Hence, a node protection bypass tunnel will be available via the protector to the proxy node, but actually be terminated by the protector.

o The second category require a PLR to have knowledge of the PW egress endpoint protection schema. The primary PE advertises the context identifier as a regular IP address, while the protector advertises it by using an explicit "context identifier" object, which MUST be understood by the PLR. The "context identifier" object requires an IGP extension. In both the routing domain and the TE domain, the context identifier is only reachable via the primary PE. This ensures that the transport tunnel is terminated by the primary PE. The PLR views itself as the penultimate hop of the transport tunnel, and based on the IGP "context identifier" object, it establishes or resolves a bypass tunnel to the advertiser (i.e. the protector), while avoiding the primary PE.

The mechanism in this document intends to be flexible on the approach used by a network, as long as it satisfies the above requirements for transport tunnel path and bypass tunnel path. For any approach, the coordination between a primary PE and a protector can be achieved by configuration.

4.4. Protection Models

There are two protection models based on the location of a protector. A network MAY use either model or both.

4.4.1. Co-located Protector

   In this model, the protector is a backup PE that is directly
   connected to the target CE via a backup AC, or it is a backup S-PE on
   a backup PW. That is, the protector is co-located with the backup
   (S-)PE. Examples of this model have been shown in Figure 4, Figure 5
   and Figure 6 in Section 4.2.

   In egress AC protection and egress PE node protection, when a
   protector receives traffic from the PLR, it forwards the traffic to
   the CE via the backup AC. This is shown in Figure 7, where PE2 is
   the PLR for egress AC failure, P3 is the PLR for PE2 failure, and PE4
   (the backup PE) is the protector.

```
                 |<------------- PW1 --------------->|

            - PE1 ------------- P1 ------- P3 ----- PE2 ----
           /                             PLR \     PLR      \
          /                                   \    |         \
      CE1                               bypass\    |bypass  CE2
         \                                     \   |       /
          \                                     \  |      /
           - PE3 ------------- P2 --------------- PE4 ----
                                                protector

                 |<------------- PW2 --------------->|
```

                             Figure 7

   In S-PE node protection, when a protector receives traffic from the
   PLR, it forwards the traffic over the next segment of the backup PW.
   The T-PE of the backup PW in turn forwards the traffic to the CE via
   a backup AC. This is shown in Figure 8, where P4 is the PLR for SPE1
   failure, and SPE2 (the backup S-PE) is the protector for SPE1.

```
              |<-------------- PW1 --------------->|
              |<----- SEG1 ----->|<----- SEG2 ----->|

          - TPE1 ----- P4  ----- SPE1 ------------- TPE2 -
         /             PLR \                              \
        /                   \                              \
     CE1                bypass\                             CE2
        \                     \                            /
         \                     \                          /
          - TPE3 -------------- SPE2 ------------- TPE4 -
                          protector

              |<----- SEG3 ----->|<----- SEG4 ----->|
              |<-------------- PW2 --------------->|
```
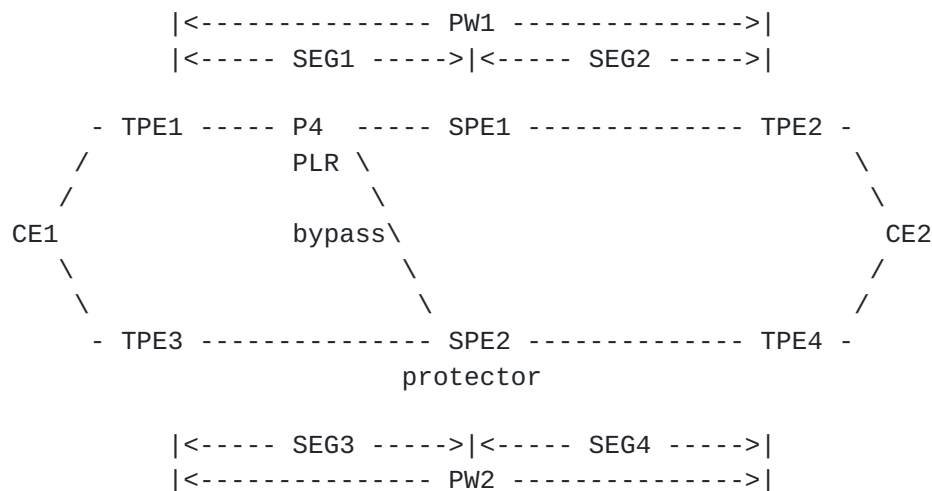
                              Figure 8

   In the co-located protector model, the number of context identifiers
   needed by a network is the number of distinct {primary PE, backup PE}
   pairs. From the perspective of scalability, the model is suitable
   for networks where the number of primary PEs and the average number
   of backup PEs per primary PE are both relatively low.

4.4.2. Centralized Protector

   In this model, the protector is a dedicated P router or PE router
   that serves the role. In egress AC protection and egress PE node
   protection, the protector may or may not be a backup PE with a direct
   connection to the target CE. In S-PE node protection, the protector
   may or may not be a backup S-PE on the backup PW.

   In egress AC protection and egress PE node protection, when the
   protector receives traffic from the PLR, if the protector has a
   direct connection (i.e. backup AC) to the CE, it forwards the traffic
   to the CE via the backup AC, which is similar to Figure 7.
   Otherwise, it forwards the traffic to a backup PE, which then
   forwards the traffic to the CE via a backup AC. This is shown in
   Figure 9, where the protector receives traffic from P3 (the PLR for
   egress PE failure) or PE2 (the PLR for egress AC failure) and
   forwards the traffic to PE4 (the backup PE). The protector may be
   protecting other PWs and other primary PEs as well, which is not
   shown in this figure for clarity.

```
                    |<------------- PW1 --------------->|

            - PE1 ------------- P1 ------- P3 ----- PE2 --
           /                             PLR \     PLR   \
          /                                  \    /       \
         /                             bypass\   /bypass  \
        /                                     \ /          \
     CE1                                    protector       CE2
        \                                      \           /
         \                                      \         /
          \                                      \       /
           \                                      \     /
            - PE3 ------------- P2 ----------------PE4 --

                    |<------------- PW2 --------------->|
```

                              Figure 9

   In S-PE node protection, when the protector receives traffic from the
   PLR, if the protector is a backup S-PE of the backup PW, it forwards
   the traffic over the next segment of the backup PW, and the T-PE of
   the backup PW forwards the traffic to the CE via a backup AC, which
   is similar to Figure 8. Otherwise, the protector first forwards the
   traffic to the backup S-PE, which then forwards the traffic over the
   next segment of the backup PW. Finally, the T-PE of the backup PW
   forwards the traffic to the CE via a backup AC. This is shown in
   Figure 10, where the protector forwards traffic to SPE2 (the backup
   S-PE), SPE2 forwards the traffic to TPE4 via SEG4, and TPE4 finally
   forwards traffic to CE2. The protector may be protecting other PW
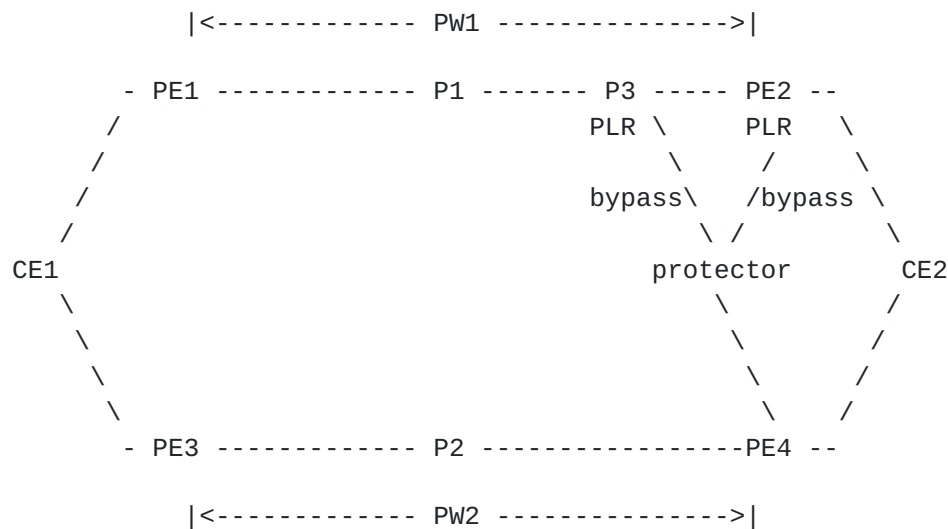   segments and other primary S-PEs as well, which is not shown in this
   figure for clarity.

```
             |<-------------- PW1 --------------->|
             |<----- SEG1 ----->|<----- SEG2 ----->|

          - TPE1 ----- P4  ----- SPE1 ------------- TPE2 -
           /            PLR \                           \
          /                  \                           \
         /            bypass\                             \
        /                    \                             \
     CE1                   protector                        CE2
        \                     \                            /
         \                     \                          /
          \                     \                        /
           \                     \                      /
           - TPE3 -------------- SPE2 -------------- TPE4 -

             |<----- SEG3 ----->|<----- SEG4 ----->|
             |<-------------- PW2 --------------->|
```
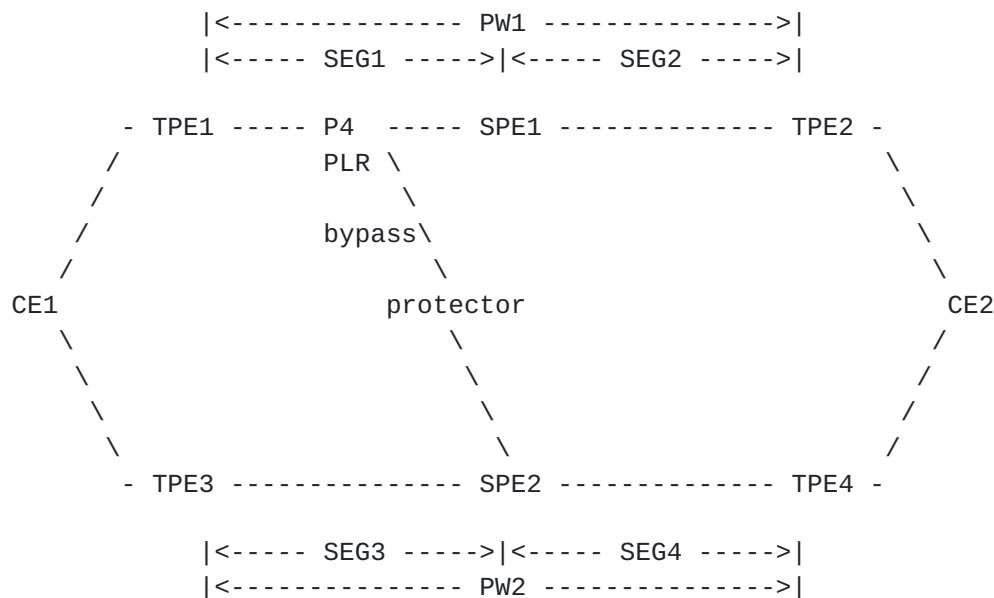
                          Figure 10

   The centralized protector model allows multiple primary PEs to share
   one protector. Each primary PE may need only one protector.
   Therefore, the number of context identifiers needed by a network may
   be bound to the number of primary PEs.

4.5. Transport Tunnel

   A PW is associated with a pair of {primary PE, protector}, which is
   represented by a unique context identifier. The ingress PE of the PW
   sets up or resolves a transport tunnel by using the context
   identifier rather than a private IP address of the primary PE as
   destination. This not only ensures that the PW is transported to the
   primary PE, but also facilitates bypass tunnel establishment at PLR,
   because the context identifier contains the identity of the protector
   as well. This is also the case for a multi-segment PW, where the
   ingress PE and egress PE are T/S-PEs.

   An ingress PE learns the association between a PW and a context
   identifier from the primary PE, which MUST advertise the context
   identifier as a "third party next hop" via the IPv4/v6 Interface_ID
   TLV [RFC3471, RFC3472] in the LDP Label Mapping message of the PW.

4.6. Bypass Tunnel

   A PLR may protect multiple PWs associated with one or multiple pairs
   of {primary PE, protector}. The PLR MUST establish a bypass tunnel to
   each protector for each context identifier associated with that
   protector. The destination of the bypass tunnel MUST be the context

identifier (Section 4.3.1). Since the PLR is a transit router of the
transport tunnel, it SHOULD derive the context identifier from the
destination of the transport tunnel.

For examples, in Figure 7 and Figure 9, a bypass tunnel is
established from PE2 (PLR for egress AC failure) to the protector,
and another bypass tunnel is established from P3 (PLR for egress node
failure) to the protector. In Figure 8 and Figure 10, a bypass
tunnel is established from P4 (PLR for S-PE failure) to the
protector.

In local repair, a PLR reroutes traffic to the protector through a
bypass tunnel, with PW label intact in the packets. This normally
involves pushing a label to the label stack, if the bypass tunnel is
an MPLS tunnel, or pushing an IP header to the packets, if the bypass
tunnel is an IP tunnel. Upon receipt of the packets, the protector
forwards them based on the PW label. Specifically, the protector
uses the bypass tunnel as a context to determine the primary PE's
label space. If the bypass tunnel is an MPLS tunnel, the protector
should have assigned a non-reserved label to the bypass tunnel, and
hence this label can serve as the context. If the bypass tunnel is
an IP tunnel, the context identifier should be the destination
address of IP header.

To be useful for local repair, a bypass tunnel MUST have the property
that it is not affected by any topology changes caused by the
failure. It should remain effective during local repair, until the
traffic is moved to another fully functional path, i.e. either the
same PW over a fully functional transport tunnel, or another fully
functional PW.

4.7. Forwarding State on Protector

A protector learns PW labels from all the primary PEs that it
protects (Section 6.2), and maintains the PW labels in separate label
spaces on a per primary PE basis. In the control plane, each label
space is identified by the context identifier of the corresponding
{primary PE, protector}. In the forwarding plane, it is indicated by
the bypass tunnel(s) destined for the context identifier.

4.7.1. Examples of Co-located Protector

In Figure 7, PE4 is a co-located protector that protects PW1 against
egress AC failure and egress node failure. It maintains a label
space for PE2, which is identified by the context identifier of {PE2,
PE4}. It learns PW1's label from PE2, and installs an forwarding
entry for the label in that label space. The nexthop of the

forwarding entry indicates a label pop with outgoing interface
pointing to the backup AC PE4-CE2.

e In Figure 8, SPE2 is a co-located protector that protects PW1
against S-PE failure. It maintains a label space for SPE1, which is
identified by the context identifier of {SPE1, SPE2}. It learns
SEG1's label from SPE1, and installs a forwarding entry in the label
space. The nexthop of the forwarding entry indicates a label swap to
SEG4's label.

4.7.2. Examples of Centralized Protector

In the centralized protector model, for each primary PW of which the
protector is not a backup (S-)PE, the protector MUST also learn the
label of the backup PW from the backup (S-)PE (Section 6.3). This is
the backup (S-)PE that the protector will forward traffic to. The
protector MUST install a forwarding entry with label swap from the
primary PW's label to the backup PW's label.

In Figure 9, the protector is a centralized protector that protects
PW1 against egress AC failure and egress node failure. It maintains
a label space for PE2, which is identified by the context identifier
of {PE2, protector}. It learns PW1's label from PE2, and PW2's label
from PE4. It installs a forwarding entry for PW1's label in the
label space. The nexthop of the forwarding entry indicates a label
swap to PW2's label.

In Figure 10, the protector is a centralized protector that protects
the PW segment SEG1 of PW1 against the node failure of SPE1. It
maintains a label space for SPE1, which is identified by the context
identifier of {SPE1, protector}. It learns SEG1's label from SPE1,
and learns SEG3's label from SPE2. It installs a forwarding entry
for SEG1's label in the label space. The nexthop of the forwarding
entry indicates a label swap to SEG3's label.

5. Revertive Behavior

Subsequent to local repair, there are three strategies for a network
to restore traffic to a fully functional path.

o Global revertive mode

   If the ingress CE is multi-homed (Figure 1), it MAY switch the
   traffic to the backup AC which is bound to the backup PW.
   Alternatively, if the ingress PE hosts a backup PW (Figure 2), the
   ingress PE MAY switch the traffic to the backup PW. These
   procedures are referred to as global repair. Possible triggers of

global repair include PW status notification, VCCV, BFD, end-to-
end OAM between CEs, etc.

o Control plane revertive mode

   In egress PE node protection and S-PE node protection, it is
   possible that the failure is limited to the link between the PLR
   and the primary PE, whereas the primary PE is still operational.
   In this case, the PLR or an upstream router on the transport
   tunnel MAY reroute the tunnel around the link via an alternative
   path to the primary PE. Thus, the transport tunnel can heal and
   continue to carry the PW to the primary PE. This procedure is
   driven by control plane convergence on the new topology, and is
   referred to as control plane repair.

o Local revertive mode

   The PLR MAY move traffic back to the primary PW, after the failure
   is resolved. In egress AC protection, upon detecting that the
   primary AC is restored, the PLR MAY start forwarding traffic over
   the AC again. Likewise, in egress PE node protection and S-PE
   node protection, upon detecting that the primary PE is restored,
   the PLR MAY re-establish the transport tunnel to the primary PE,
   and move the traffic from the bypass tunnel back to the transport
   tunnel. These procedures are referred to as local reversion.

It is RECOMMENDED that the fast protection mechanism SHOULD be used
in conjunction with the global revertive mode. Particularly in the
case of egress PE and S-PE node failures, if the ingress PE or the
protector loses communication with the (S-)PE for an extensive period
of time, LDP session may go down. Consequently, the ingress PE may
bring down the primary PW completely, or the protector may remove the
forwarding entry of the primary PW label. In either case, the
service will be disrupted. In other words, although the mechanism
can temporarily repair traffic, control plane state may eventually
expire if the failure persists. Therefore, the global revertive mode
SHOULD take place in a timely manner to move traffic to a fully
functional path.

The control plane revertive mode may automatically happen as part of
the convergence of control plane protocols. However, it is only
applicable to the specific link failure scenario described above.

The local revertive mode is optional. In the circumstances where the
failure is caused by resource flapping, local reversion MAY be
dampened to limit potential disruption. Local revertive mode MAY be
disabled completely by configuration.

6. **LDP Extensions**

   As described in previous sections, a targeted LDP session MUST be
   established between each pair of primary PE and protector. The
   primary PE sends Label Mapping message over this session to advertise
   primary PW labels to the protector. In the centralized protector
   model, a targeted LDP session MUST also be established between a
   backup (S-)PE and the protector. The backup PE sends Label Mapping
   message over this session to advertise backup PW labels to the
   protector.

   To facilitate the procedures, this document defines a new "Protection
   FEC Element" TLV. The Label Mapping messages of both the LDP
   sessions above MUST carry this TLV to identify a primary PW.
   Specifically, in the centralized protector model, the Protection FEC
   Element TLV advertised by a backup (S-)PE MUST match the one
   advertised by the primary PE, so that the protector can associate the
   primary PW's label with the backup PW's label, and perform a label
   swap. The backup (S-)PE builds such a Protection FEC Element TLV
   based on local configuration.

   This document also defines the encoding of Capability Parameter TLV
   [RFC5561] for a new "Egress Protection Capability", to allow a
   protector to announce its capability of processing the above
   Protection FEC Element TLV and performing context specific label
   switching for PW labels.

   The procedures in this section are only applicable, if the protector
   advertises the Egress Protection Capability, the primary PE supports
   the advertisement of the Protection FEC Element TLV, and in the
   centralized protector model, the backup PE also supports the
   advertisement of the Protection FEC Element TLV.

6.1. **Egress Protection Capability TLV**

   A protector MUST advertise the Egress Protection Capability TLV in
   its Initialization message and Capability message, over the LDP
   session with a primary PE. In the centralized protector model, the
   protector MUST also advertise the TLV over the LDP session with a
   backup PE. The TLV carries one or multiple context identifiers. To
   the primary PE, the TLV MUST carry the context identifier of the
   {primary PE, protector}. In the centralized protector model, the TLV
   MUST carry to the backup PE multiple context identifiers, one for
   each {primary PE, protector} where the backup PE serves as a backup
   for the primary PE. This TLV MUST NOT be advertised by the primary
   PE or the backup PE to the protector.

The processing of the Egress Protection Capability TLV by a receiving
router MUST follow the procedures defined in RFC5561. In particular,
the router MUST advertise PW information to the protector by using
the Protection FEC Element TLV, only after it has received the Egress
Protection Capability TLV from the protector. It MUST validate each
context identifier included in the TLV, and advertise the information
of only the PWs that are associated with the context identifier. It
MUST withdraw previously advertised Protection FEC TLVs, when the
protector has withdrawn a previously advertised context identifier or
the entire Egress Protection Capability TLV via Capability message.

The encoding of the Egress Protection Capability TLV is defined as
below. It conforms to the format of Capability Parameter TLV
specified in RFC5561.

```
  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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |U|F|  Egress Protection (TBD)  |             Length            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |S| Reserved    |                                               |
 +-+-+-+-+-+-+-+-+                                               |
 |                                                               |
 ~               Capability Data = context identifier(s)         ~
 |                                                               |
 |                                               +-+-+-+-+-+-+-+-+
 |                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
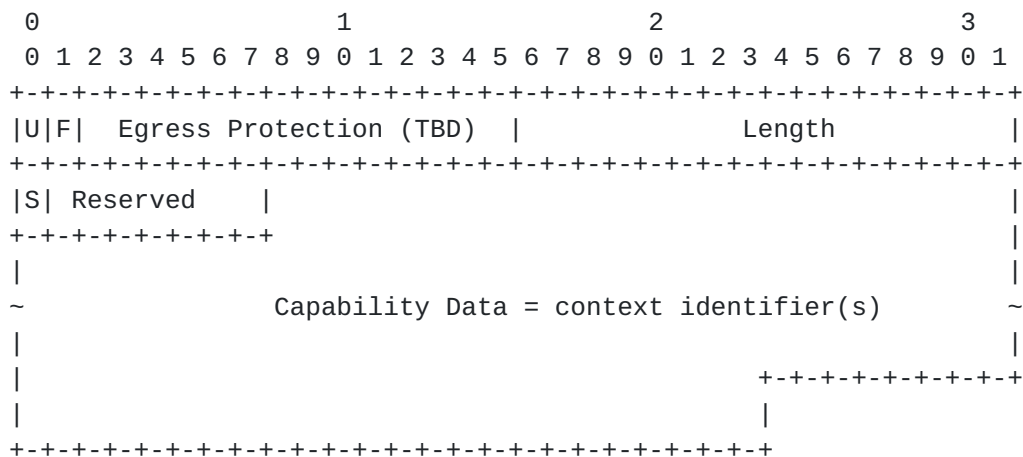
                            Figure 11

The U-bit MUST be set to 1 so that a receiver MUST silently ignore
this TLV if unknown to it, and continue processing the rest of the
message.

The F-bit MUST be set to 0 since this TLV is sent only in
Initialization and Capability messages, which are not forwarded.

The TLV Code Point is TBD. It needs to be assigned by IANA.

The S-bit indicates whether the sender is advertising (S=1) or
withdrawing (S=0) the capability.

The "Capability Data" is encoded with the context identifier of the
{primary PE, protector}.

6.2. **PW Label Distribution from Primary PE to Protector**

   A primary PE MUST advertise a primary PW's label to a protector by
   sending a Label Mapping message. The message includes a Protection
   FEC Element TLV (see Section 6.4 for encoding), and an Upstream-
   Assigned Label TLV [RFC6389] encoded with the PW's label. The
   combination of the Protection FEC Element TLV and the PW label
   represents the primary PE's forwarding state for the PW. The Label
   Mapping message MUST also carry an IPv4/v6 Interface_ID TLV [RFC6389,
   RFC3471] encoded with the context identifier of the {primary PE,
   protector}.

   The protector that receives this Label Mapping message MUST install a
   forwarding entry for the PW label in the label space identified by
   the context identifier. The nexthop of the forwarding entry MUST
   ensure packets to be sent towards the target CE via a backup AC or a
   backup (S-)PE, depending on the protection scenario. The protector
   MUST silently discard a Label Mapping message if the included context
   identifier is unknown to it.

6.3. **PW Label Distribution from Backup PE to Protector**

   In the centralized protector model, a backup PE MUST advertise a
   backup PW's label to the protector by sending a Label Mapping
   message. The message includes a Protection FEC Element TLV and a
   Generic Label TLV encoded with the backup PW's label. This
   Protection FEC Element MUST be identical to the Protection FEC
   Element TLV that the primary PE advertises to the protector
   (Section 6.2). This is achieved through configuration on the backup
   PE. The context identifier MUST NOT be encoded in Interface_ID TLV
   in this message.

   The protector that receives this Label Mapping message MUST associate
   the backup PW with the primary PW, based on the common Protection FEC
   Element TLV. It MUST distinguish between the Label Mapping message
   from the primary PE and the Label Mapping message from the backup PE
   based on the respective presence and absence of context identifier in
   Interface_ID TLV. It MUST install a forwarding entry for the primary
   PW's label in the label space identified by the context identifier.
   The nexthop of the forwarding entry MUST indicate a label swap to the
   backup PW's label, followed by a label push or IP header push for a
   transport tunnel to the backup PE.

6.4. **Protection FEC Element TLV**

   The Protection FEC Element TLV has type 0x83. Its format is defined
   as below:

```
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  Type(0x83)  |    Reserved   | Encoding Type |    Length     |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    |                                                               |
    ~                      PW Information                           ~
    |                                                               |
    |                               +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                            Figure 12

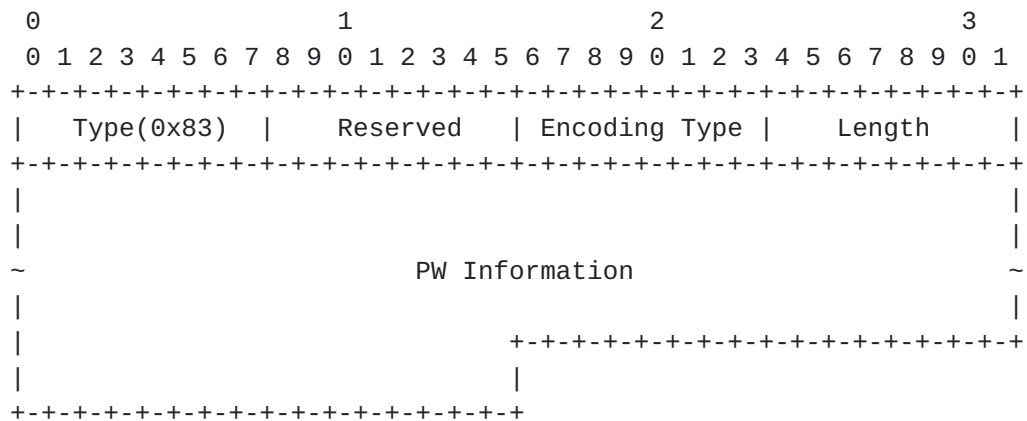   - Encoding Type

       Type of format that PW Information field is encoded.

   - Length

       Length of PW Information field in octets.

   - PW Information

       Field of variable length that specifies a PW

   For Encoding Type, 1 is defined for the PWid FEC Element format, and
   2 is defined for the Generalized PWid FEC Element format [RFC4447].

6.4.1. Encoding Format for PWid

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Type(0x83)  |   Reserved   |  Enc Type(1)  |  Length(16)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Ingress PE Address                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Egress PE Address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Group ID                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         PW ID                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C|         PW Type          |              Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                          Figure 13

- Ingress PE Address

    IP address of the ingress PE of PW.

- Egress PE Address

    IP address of the egress PE of PW.

- Group ID

    An arbitrary 32-bit value that represents a group of PWs and that
    is used to create groups in the PW space.

- PW ID

    A non-zero 32-bit connection ID that, together with the PW Type
    field, identifies a particular PW.

- Control word bit (C)

    A bit that flags the presence of a control word on this PW. If C
    = 1, control word is present; If C = 0, control word is not
    present.

- PW Type

    A 15-bit quantity that represents the type of PW.

6.4.2. Encoding Format for Generalized PWid

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type(0x83)  |    Reserved   |  Enc Type(2)  |    Length     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Ingress PE Address                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Egress PE Address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C|         PW Type         |               Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   AGI Type    |    Length     |         Value                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                   AGI  Value (contd.)                         ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   AII Type    |    Length     |         Value                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                   SAII  Value (contd.)                        ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   AII Type    |    Length     |         Value                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                   TAII Value (contd.)                         ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
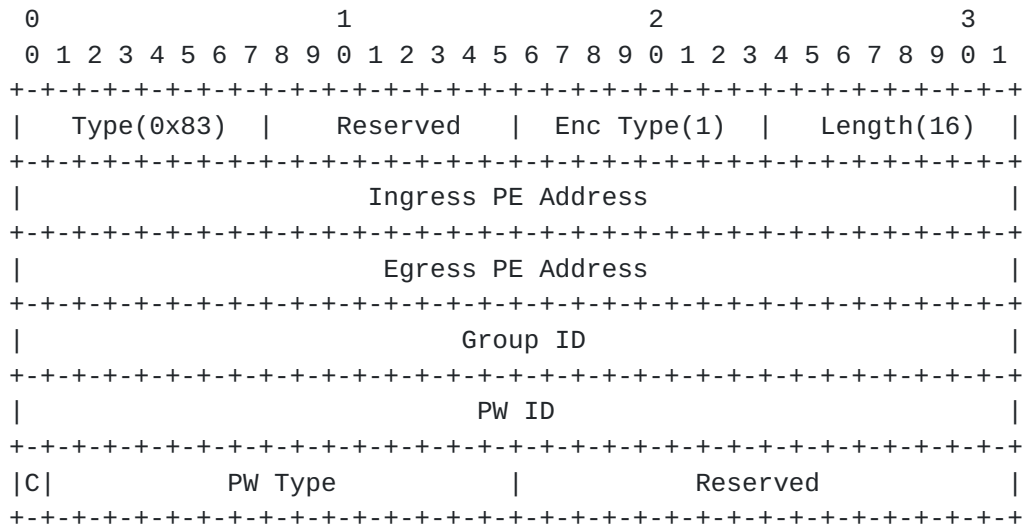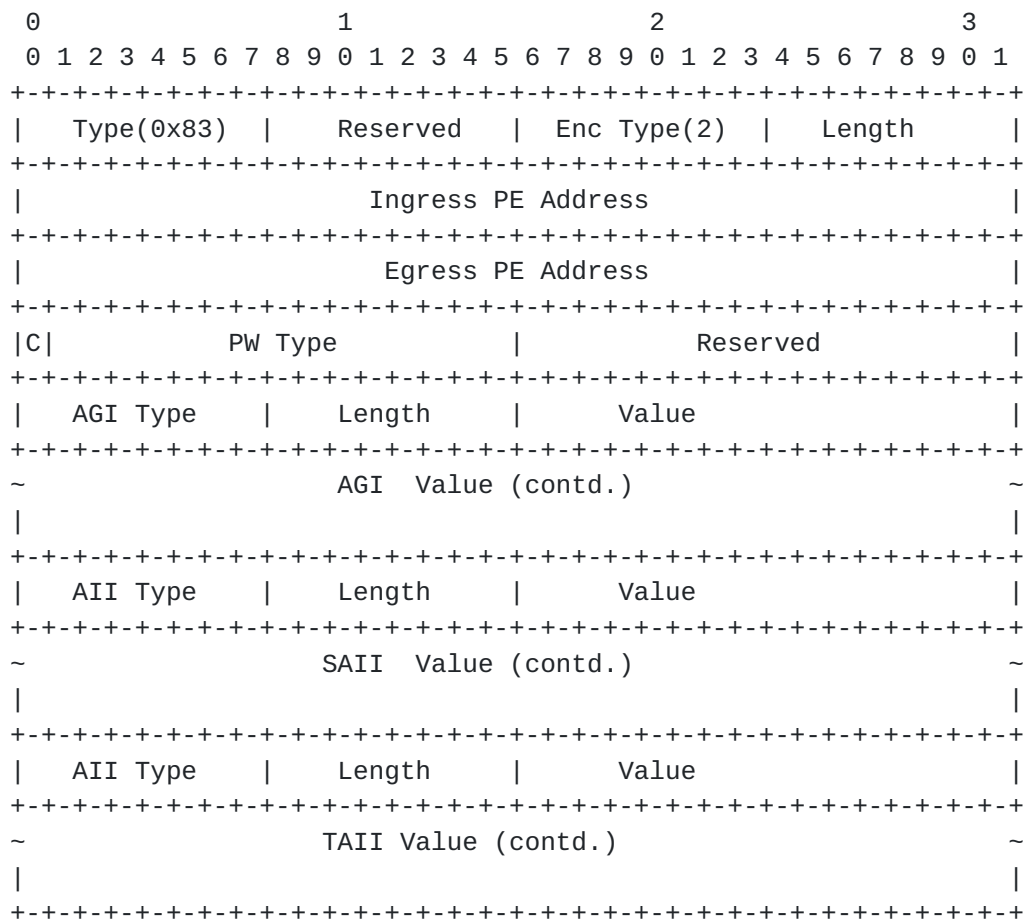
Figure 14

  - Ingress PE Address

    IP address of the ingress PE of PW.

  - Egress PE Address

    IP address of the egress PE of PW.

  - Control word bit (C)

    A bit that flags the presence of a control word on this PW. If C
    = 1, control word is present; If C = 0, control word is not
    present.

  - PW Type

    A 15-bit quantity that represents the type of PW.

   - AGI Type, Length, Value, AGI Value

      Attachment Group Identifier of PW.

   - SAII Type, Length, Value, SAII Value

      Source Attachment Individual Identifier of PW.

   - TAII Type, Length, Value, TAII Value

      Target Attachment Individual Identifier of PW.

7. IANA Considerations

   This document defines the encoding of the Capability Parameter TLV
   for the new "Egress Protection Capability" in Section 6. This would
   require IANA to assign a TLV Code Point to it.

   This document defines a new LDP Protection FEC Element TLV in
   Section 6. IANA has assigned the type value 0x83 to it.

```
   Value  Hex   Name                    Label Advertisement Discipline
   -------------------------------------------------------------------
   131    0x83  Protection FEC Element  DU
```

8. Security Considerations

   The security considerations discussed in RFC5036, RFC5331, RFC3209,
   and RFC4090 apply to this document. There is no additional
   consideration.

9. Acknowledgements

   This document leverages work done by Hannes Gredler, Yakov Rekhter,
   Minto Jeyananth, Kevin Wang and several on MPLS edge protection.
   Thanks to Nischal Sheth and Bhupesh Kothari for their contribution.
   Thanks to John E Drake, Andrew G Malis, Alexander Vainshtein, Steward
   Bryant, and Mach Chen for valuable comments that helped shape this
   document and improve its clarity.

10. References

10.1. Normative References

   [RFC3985] Bryant, S., Ed. and P. Pate, Ed., "Pseudo Wire Emulation
              Edge-to-Edge (PWE3) Architecture", RFC 3985,
              DOI 10.17487/RFC3985, March 2005,
              <http://www.rfc-editor.org/info/rfc3985>.

   [RFC5659] Bocci, M. and S. Bryant, "An Architecture for Multi-
              Segment Pseudowire Emulation Edge-to-Edge", RFC 5659,
              DOI 10.17487/RFC5659, October 2009,
              <http://www.rfc-editor.org/info/rfc5659>.

   [RFC4447] Martini, L., Ed., Rosen, E., El-Aawar, N., Smith, T., and
              G. Heron, "Pseudowire Setup and Maintenance Using the
              Label Distribution Protocol (LDP)", RFC 4447,
              DOI 10.17487/RFC4447, April 2006,
              <http://www.rfc-editor.org/info/rfc4447>.

   [RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream
              Label Assignment and Context-Specific Label Space",
              RFC 5331, DOI 10.17487/RFC5331, August 2008,
              <http://www.rfc-editor.org/info/rfc5331>.

   [RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed.,
              "LDP Specification", RFC 5036, DOI 10.17487/RFC5036,
              October 2007, <http://www.rfc-editor.org/info/rfc5036>.

   [RFC5561] Thomas, B., Raza, K., Aggarwal, S., Aggarwal, R., and JL.
              Le Roux, "LDP Capabilities", RFC 5561,
              DOI 10.17487/RFC5561, July 2009,
              <http://www.rfc-editor.org/info/rfc5561>.

   [RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S.
              Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1
              Functional Specification", RFC 2205, DOI 10.17487/RFC2205,
              September 1997, <http://www.rfc-editor.org/info/rfc2205>.

   [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
              and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
              Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001,
              <http://www.rfc-editor.org/info/rfc3209>.

   [RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast
              Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090,
              DOI 10.17487/RFC4090, May 2005,
              <http://www.rfc-editor.org/info/rfc4090>.

   [RFC5286] Atlas, A., Ed. and A. Zinin, Ed., "Basic Specification for
              IP Fast Reroute: Loop-Free Alternates", RFC 5286,
              DOI 10.17487/RFC5286, September 2008,
              <http://www.rfc-editor.org/info/rfc5286>.

   [RFC5714] Shand, M. and S. Bryant, "IP Fast Reroute Framework",
              RFC 5714, DOI 10.17487/RFC5714, January 2010,
              <http://www.rfc-editor.org/info/rfc5714>.

[RFC3471]   Berger, L., Ed., "Generalized Multi-Protocol Label
            Switching (GMPLS) Signaling Functional Description",
            RFC 3471, DOI 10.17487/RFC3471, January 2003,
            <http://www.rfc-editor.org/info/rfc3471>.

[RFC3472]   Ashwood-Smith, P., Ed. and L. Berger, Ed., "Generalized
            Multi-Protocol Label Switching (GMPLS) Signaling
            Constraint-based Routed Label Distribution Protocol (CR-
            LDP) Extensions", RFC 3472, DOI 10.17487/RFC3472, January
            2003, <http://www.rfc-editor.org/info/rfc3472>.

[RFC3031]   Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol
            Label Switching Architecture", RFC 3031,
            DOI 10.17487/RFC3031, January 2001,
            <http://www.rfc-editor.org/info/rfc3031>.

[RFC2328]   Moy, J., "OSPF Version 2", STD 54, RFC 2328,
            DOI 10.17487/RFC2328, April 1998,
            <http://www.rfc-editor.org/info/rfc2328>.

[RFC5880]   Katz, D. and D. Ward, "Bidirectional Forwarding Detection
            (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010,
            <http://www.rfc-editor.org/info/rfc5880>.

[RFC6389]   Aggarwal, R. and JL. Le Roux, "MPLS Upstream Label
            Assignment for LDP", RFC 6389, DOI 10.17487/RFC6389,
            November 2011, <http://www.rfc-editor.org/info/rfc6389>.

[RFC6391]   Bryant, S., Ed., Filsfils, C., Drafz, U., Kompella, V.,
            Regan, J., and S. Amante, "Flow-Aware Transport of
            Pseudowires over an MPLS Packet Switched Network",
            RFC 6391, DOI 10.17487/RFC6391, November 2011,
            <http://www.rfc-editor.org/info/rfc6391>.

[IP-LDP-FRR-MRT]
            Atlas, A. and R. Kebler, "An Architecture for IP/LDP Fast-
            Reroute Using Maximally Redundant Trees", draft-ietf-
            rtgwg-mrt-frr-architecture (work in progress), 2011.

10.2. Informative References

[RFC5920]   Fang, L., Ed., "Security Framework for MPLS and GMPLS
            Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010,
            <http://www.rfc-editor.org/info/rfc5920>.

Authors' Addresses

   Yimin Shen
   Juniper Networks
   10 Technology Park Drive
   Westford, MA 01886
   USA

   Phone: +1 9785890722
   Email: yshen@juniper.net

   Rahul Aggarwal
   Arktan, Inc

   Email: raggarwa_1@yahoo.com

   Wim Henderickx
   Alcatel-Lucent
   Copernicuslaan 50
   2018 Antwerp
   Belgium

   Email: wim.henderickx@alcatel-lucent.be

   Yuanlong Jiang
   Huawei Technologies

   Email: jiangyuanlong@huawei.com