Entropy label for seamless MPLS
draft-ravisingh-mpls-el-for-seamless-mpls-00

Abstract

   This document describes how entropy labels can be used for load
   balancing in a seamless MPLS architecture. The definition of the
   control plane and data plane behavior at LSP stitching points; and at
   the ingress of an LSP in a hierarchy of LSPs, as described in this
   document, brings the benefits of entropy labels to seamless MPLS as
   MPLS deployments proliferate in the access and aggregation networks.

   This document updates RFC 6790.

Status of this Memo

Copyright and License Notice

Table of Contents

1 Introduction

[EL-RFC] specifies a way to implement load-balancing in an MPLS network such that sub-flows of an LSP may be identified and sent on different paths through the network. This is achieved by using entropy labels (ELs) to abstract out the flow-identifying information into the entropy label and inserting the entropy label underneath the LSP label. The transit LSRs perform the load-balancing hash-computation, on the label-stack alone, to effect a good load-balancing outcome without a need to parse inner headers.

The key feature of [EL-RFC] is that it defines the EL in the context of a given LSP. [EL-RFC] defines the signaling extensions by which entropy label capability might be signaled for LSPs setup by RSVP-TE, LDP or [LU-BGP]. While that works well for individual LSPs, there are additional issues to consider for the seamless MPLS architecture [S-MPLS].

The currently-under-definition framework for seamless MPLS proposes an architecture ([S-MPLS]) that shall enable the setting-up of MPLS LSPs from access nodes to access nodes using a medley of signaling protocols and statically configured LSPs. There are special EL-related considerations that need to be dealt with to make EL more suitable for seamless MPLS.

This document defines additional abstractions and rules for the use of entropy-label with LSP stitching/hierarchy to enable the use of ELs for the seamless MPLS architecture. This document describes how entropy labels may be used when the LSP has been setup by stitching LSP segments or by tunneling the LSP over other LSPs. It is conceivable that different signaling protocols are in use to create an e2e LSP.

LSP stitching is the process of connecting LSP segments in the data plane to form a single e2e data plane LSP. This is achieved by setting up LSP segments through signaling or through management action, and then signaling an e2e LSP that "uses" these LSP segments as hops in its path. The procedures for LSP stitching are described in [STITCHING]. Labeled data traffic flowing over e2e MPLS LSPs, that have been setup using multiple different protocols (by stitching together segments), would benefit from having the entropy label be included in it.

LSP hierarchy is defined in [MPLS-ARCH] and [GMPLS-HIER]. Usage of entropy label in LSP hierarchies has some peculiar practical issues that will benefit from some additional flexibility in inserting ELs for a specific layer in an LSP hierarchy.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
document are to be interpreted as described in RFC 2119 [RFC2119].

The following acronyms/terms are used:

e2e: End to end LSP that has been setup by stitching together LSP
     segments

ECMP: Equal Cost Multi-Path

EL: Entropy Label

ELC: Entropy Label Capability or Entropy Label Capable

ELI: Entropy Label Indicator

Intrinsic ELC: Entropy label capability/capable as in [EL-RFC]. In
               this document, an LSP is considered to be
               "intrinsically" EL-capable when the:
               - the ingress of the LSP has the ability to compute
                 and PUSH the EL before PUSHing the ELI before
                 PUSHing the LSP label; and
               - the egress/PHR of the LSP-segment has the ability
                 to POP the (ELI+EL) at the egress/PHR while
                 POPing-transport-label/ELI-is-top-label
                 respectively.

LAG: Link Aggregation Group

LER: Label Edge Router

LSP: Label Switched Path

LSR: Label Switching Router

Notional ingress: Ingress LER for an LSP segment that is inserting
                  the (ELI+EL) on data traffic going over an e2e LSP

Notional egress: egress LER for an LSP segment that is removing the
                 (ELI+EL) from data traffic going over an e2e LSP

Notional LSP segment: the portion of the e2e LSP between a notional
                      ingress and a notional egress

PHP: Penultimate Hop Popping

PHR: Penultimate Hop Router

UHP: Ultimate Hop Popping

NOTE: this document references the (ELI+EL) pair simply as EL when
      the presence of the ELI is of no significance for the issue
      being described. The presence of ELI is mandatory as per
      [EL-RFC] when EL is in use.

3. Key attributes of the entropy label solution: Summary from [EL-RFC]
     - Transport-label-PUSHing router inserts (ELI+EL)
       The (ELI+EL) insertion is done, if at all, by a router that is
       PUSHing the transport LSP's label.

     - Ingress-LER (transport-label-PUSHing-router) inserts (ELI+EL)
       only if the PHR/egress has signaled ability to strip it off.

     - Transport-label-POPing router POPs (ELI+EL) PHR/egress of the
       LSP is responsible for POPing off the (ELI+EL) after it has
       been exposed as the top label on the packet due to POPing the
       transport label. The removal of the (ELI+EL) is done either
       when the ELI is the top label; or when the ELI is next label
       below the top label being POPed.

     - Max-payload size for the LSP gets reduced by 8 bytes after the
       insertion of the (ELI+EL).

4. Problems and Motivation

   [EL-RFC] defines EL signaling/usage suitable for single-segment LSPs.
   However, as MPLS proliferates in the network access leading to the
   setup of e2e LSPs using LSP stitching and hierarchies, there is a
   need to define the EL behavior for LSP stitching and LSP hierarchies.

   [EL-RFC] does not explicitly specify the EL-signaling-interaction
   between stitched LSPs segments. Similarly, peculiarities in the data-
   plane related to LSP stitching need further specification. Likewise,
   the signaling and data-plane peculiarities for using EL over LSP
   hierarchies could be further specified.

   It is desirable to get the benefits of EL even for stitched LSPs.

   Certain aspects peculiar to stitched LSPs need additional handling to
   increase the applicability of [EL-RFC]. [EL-RFC] needs to be extended

to define the behavior for LSP stitching and LSP hierarchies
(tunneling) when using EL.

The sub-sections below list the specific additional requirements for
making entropy label more deployable when used with LSP stitching,
and LSP hierarchy.

4.1 EL applicability for seamless MPLS

The seamless MPLS architecture relies on the use of LSP stitching and
hierarchy to signal an e2e LSP between access-nodes, such that the
e2e LSP is going over aggregation/transport/cores nodes.

The signaling of such e2e LSPs is enabled by using the
stitching/hierarchy mechanisms that exist, using [LU-BGP]/LDP/RSVP.

The rules of section 5 provide a general-purpose way for the use of
ELs across e2e LSPs by defining:

- the rules of ELC propagation at stitching points;

- the data-plane guidelines for the stitching point LSR; and

- the data-plane guidelines for LSP hierarchies for inserting
  (ELI+EL) at ingress LER of an LSP in an LSP hierarchy.

4.2 EL for LSP stitching

A stitched e2e LSP might be stitched from greater than 2 segment LSPs
(along the length of the e2e LSP), with 2 LSPs forming the stitch at
each stitching point.

An LSP segment is considered to be intrinsically EL capable when it
supports the attributes summarized in section 3.

Some of the segment LSPs in the e2e LSP may intrinsically support EL
and some may not. So, the e2e LSP may not intrinsically support EL
from end to end. However, to obtain the benefits of EL for stitched
LSPs, it is important that an EL should be present on the data
packets traversing as many segments of the e2e LSP as is possible
within data plane abilities of the routers on the way.

In using EL with LSP stitching, the aims are BOTH of the following:
     a. Get EL benefits wherever possible: on all segments where
        possible. Just because a given segment does not support EL
        is not a reason to deny EL benefits to other segments of the
        e2e LSP.

b. Not run into data-plane problems where an
   intermediate-segment whose ingress LER can not look deeper
   to remove EL when the subsequent segment does not support EL.

- Independent setup of LSP segments:

LSP stitching typically relies on LSP segments that have been
independently setup. In an e2e LSP (made of stitched segments), it is
unlikely that all of the stitching points (i.e., segment LSP end
points) as well as the ultimate ingress and ultimate egress support
EL as defined in section 3.

However, there would be individual LSP segments that would completely
satisfy the requirements of section 2 (i.e. are intrinsically EL
capable). This document describes how EL may be used for (portions
of) the e2e LSP while still working within the framework for [EL-
RFC].

     S---A---B---C---D

In the above topology, for an e2e LSP from S to D, the segments AB
and CD could be intrinsically EL capable while the segments SA, & BC
may not be. For data traffic going over the LSP from S to D, using EL
on the segments AB and CD would be beneficial for load-balancing over
LAGs/ECMP.

- Dealing with different protocols being used to setup the segments
  of the e2e LSP.

4.2.1 Spectrum of EL usage behaviors required to be supported for stitched LSPs

To cater for an incremental deployment of intrinsically-ELC routers
in a network, the multiple different modes for EL use with LSP
stitching need to be to be supported.

The spectrum of supported behaviors are listed below by referencing
the following diagram.

         S1   S2   S3   S4

   A-----------B-----------C-----------D-----------E

LSP segments S1, S2, S3, S4 are between LERs A/B/C/D/E. There may or
may not be other routers between the per-segment ingress<->egress
LERs.

Transport LSP signaling protocol: could be any: LDP/RSVP/([LU-BGP]
tunneled over RSVP/LDP).

4.2.1.1 Entropy label for per-segment LSP

Each of the segments will have their independent EL capability based
on BOTH the:

    - Per-segment ingress having the ability to insert the EL.

    - Per-segment egress (or PH router) having the ability to strip
       the EL.

This is very similar to [EL-RFC] with the additional data-plane rule
of section 5.2.2.1 "A. Rationalizing EL for the outgoing LSP
segment:".

Reasoning for why per-segment EL may be attractive for certain use
scenarios:

Opportunity to get benefits on those segments where EL benefits are
available. Even though the e2e LSP may not support ELC, this allows
the EL benefits on those segments that are EL-capable.

4.2.1.2 Entropy label for notional-segment-LSP(s)

In the case of stitched LSPs, it is useful to:

    - Insert EL at first per-segment ingress LER (per-segment ingress
      LER closest to the e2e ingress LER) that has the ability to
      insert EL.
    - Carry the EL on the data packets as far along the stitched LSP as
      the last per-segment egress LER that ability to strip the EL
      on a series of contiguous EL-supporting segments.

The above is enabled by the rules of section "5.2.1.1 Signaling ELC
at stitching points (Translation rules)".

The benefit of using EL with notional-segment LSPs:

An operator might be able to use EL for the MPLS traffic on its path
to a stitching point even though the stitching-point router (or its
PHR) itself may not have the data-plane capabilities required as in
[EL-RFC].

Additionally, even if the stitching-point (or its PHR) do have the

data-plane capabilities of [EL-RFC], it is just more efficient to
forward the data packets without having to strip the EL and then
reinsert the EL when the downstream segment is also intrinsically
ELC.

4.2.1.3 Entropy label for e2e LSP

This correspond to having the notional-LSP and the e2e LSP being the
same.

This is covered by the rules of section 5.2.1.1 "Signaling ELC at
stitching points (Translation rules):" with the additional
requirement that the data-plane be exactly the same as [EL-RFC],
i.e.
     the (ELI+EL) insertion is done by a label PUSHing router,
     the (ELI+EL) POP is done by the PHR/egress for the e2e LSP.

4.3 EL for LSP hierarchy

For the purpose of highlighting the problem to be addressed and the
resultant requirements to be met, the following diagram is presented
as an example.

Let there be an LSP hierarchy with the ingress for the different
levels of LSP hierarchy being at different routers, such that each
LSP in the hierarchy is intrinsically EL capable. The individual LSPs
in the hierarchy could be a single-segment LSP or a stitched e2e LSP.

```
        S1                    D1
          \     ---------    /
           A===B===C===D===E
          /     ---------    \
        S2                    D2
```
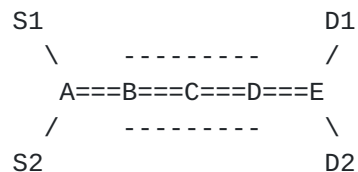
In the above topology, let there be the following LSPs:

     L1: B->D
     L2: A->E, tunneled through LSP L1
     L3: S1->D1, tunneled through LSP L2
     L4: S2->D2, tunneled through LSP L2

All of the LSPs above are assumed to be intrinsically EL capable.

4.3.1 Possibility of unnecessary reduction of max-payload of the LSP:

Even though the aim of using EL is to get better load-balancing
support, in some cases the insertion of (ELI+EL) may unnecessarily

reduce the effective payload of an LSP.

In above diagram, as per [EL-RFC] for a data packet on LSP L3, the insertion of (ELI+EL) for each of the 3 LSPs: L1, L2 and L3 is not explicitly prohibited. As a result it is possible that the packet on LSP L3, might end up with 3 (ELI+EL)s (one for each LSP level in the hierarchy) thus reducing the effective payload of the LSP L3. Likewise for L4. The presence of the (ELI+EL) for the outer LSPs L1 and L2 is not strictly useful for load-balancing the data traffic on the LSPs L3 and L4.

The solution for this issue is presented in section 5.3.2: it relies on inserting the (ELI+EL) in the context of only 1 LSP in a hierarchy.

This issues results in the following requirement for EL usage in the presence of LSP hierarchies:
- Desirability of having a single (ELI+EL) on data packets over an
  LSP hierarchy: The LSP for which the (ELI+EL) is inserted, is
  preferably the innermost intrinsically EL-capable LSP, as the
  notion of a user-flow is more specifically indicated by fields
  deeper inside the packet headers. Having an EL be present deeper
  in the packet provides load-balancing benefits of EL for the
  traversal of the packet across a longer stretch of the network.

If there is to be only 1 (ELI+EL) in the label stack, it imposes an additional data plane requirement on the ingress LER as described in section 5.3.2.

4.3.2 **Possibility of EL being non-usable for load-balancing: Even though the aim of using EL is to get better load-balancing, in some cases** the insertion of (ELI+EL) may actually offer no load-balancing benefits at all. Whether the presence of an EL offers load-balancing benefits on a given transit router, depends on:

  - whether the transit router has a LAG or an ECMP as an outgoing
    interface for the LSP, AND
  - whether the forwarding ASICs of the transit routers have the
    ability to include the EL (appearing at a specific position in
    the label stack) in the hash computation, either by:
      +   looking up the ELI and then picking the EL, or
      +   computing the hash on the maximum number of labels that it can
           pick from the label-stack for hash-computation which
           happens to also include the EL.

When the EL on a packet is outside the portion-of-the-label-stack that the ASIC of a transit router can use for hash computation, the forwarding hardware may include only the top few labels or the bottom

few labels in the hash computation. This may prevent the inclusion of
EL for hash-computation.

In the above diagram, for a data packet going over LSP L3 let the
issue of section 4.3.1 have been resolved by the router S1 inserting
the (ELI+EL) underneath the label for LSP L3 and none of the other
routers inserting the (ELI+EL). When this data packet arrives at
router C, its label stack looks thus:

```
    Label-LSP1 | Label-LSP2 | Label-LSP3 | ELI | EL
    Top-label                -> Bottom label
```

Let's say that the router C is able to include only the top 4 labels
in a label stack for the hash-computation due to the ability of its
forwarding ASICs.

So, the router C is not able to get the benefit of the presence of
the EL in the data packet. If the only ECMP/LAG in this network is
the link between C&D, then the presence of the EL serves no purpose
for the above network example and it ends up reducing the payload
capacity of the LSPs L3 and L4 by 8 bytes.

This example could be further generalized in the case of seamless
MPLS, where there may be deeper LSP hierarchies.

A transit router that has the ability to hash on an EL (based on its
depth in the label stack) does not have multiple paths; while another
router that has multiple paths and the ability hash on the EL
(appearing at a specific depth in the label stack) is unable to do so
as the EL appears outside the depth of the label stack that may be
included in the hash. In neither of the foregoing cases is the
presence of an EL helpful.

This translates into a requirement for EL: Flexibility in choice of
LSP tunnel for which EL is inserted:

There is a need to have a way by which to include an EL underneath a
specific label in a label-hierarchy based on it serving the most
useful purpose (i.e. taking into consideration location of multiple-
forwarding-paths and stack-depth-concerns).

[EL-RFC] has no way of influencing the insertion of (ELI+EL) at a
certain LSP level in the stack. Thus, there is a need for a mechanism
by which one of the many intrinsically-EL-capable LSPs in an LSP
hierarchy could be picked for inserting the (ELI+EL).

5. **EL for LSP stitching/hierarchy**
5.1 **Additional EL abstractions: specific to LSP stitching/hierarchy**

   Given the previous sections, following additional abstractions need
   to be defined to make EL more useful for LSP stitching and hierarchy.

5.2 **New abstractions: EL applicability for LSP stitching**
5.2.1 **Signaling**

   New abstractions need to be defined to handle the differences in the
   use of ELs for stitched-LSPs as compared to their use for single-
   segment LSPs.

   The differences are:

    - Notion of ingress for EL insertion:
      (ELI+EL) insertion might not necessarily be done by a
      label-PUSHing router. A stitching point where the label is being
      swapped might do the (ELI+EL) insertion, and serves as a
      "notional ingress".

    - Notion of egress for EL:
      "Notional-egress" might not be the segment egress for the segment
      of the notional-ingress.
      Even though certain stitching-points (segment LERs) might not
      support POPing (ELI+EL), it may be acceptable to let the (ELI+EL)
      continue to be on the packet since the egress of a subsequent
      segment has the capability to POP the (ELI+EL) (which may not
      necessarily be along with POPing the transport label). A
      "notional-ingress and notional-egress" pair might actually be the
      segment-ingress and segment-egress for different LSP segments that
      are part of the same e2e LSP.

   The portion of the stitched e2e LSP, between a notional-ingress and a
   notional-egress is referred to as the "notional-LSP-segment" in this
   document.

   As a packet traverses an e2e LSP, it may have an (ELI+EL) imposed on
   it and then removed at different routers.

   It is desirable for there to not be more than one instance of an
   (ELI+EL) to appear on a packet at any given time. However, the
   insertion followed by removal of an (ELI+EL) may happen more than
   once as the packet traverses the e2e LSP. Each router doing the
   (ELI+EL) insertion is the notional-ingress and each router doing the
   (ELI+EL) removal is the notional-egress (or notion EL-stripping-PH-
   router).

Thus, there may be more than 1 "notional ingress" for EL insertion,
and there may be more than 1 "notional egress" for EL removal.

For each notional "ingress ingress", there will be a "notional
egress" with the "notional ingress"es and "notional egress"es
alternating along the path of the e2e LSP when there are more than 1
notional ingress and egress for an e2e LSP.

In the simplest case, this boils down to the case of there being just
one notional ingress and one notional egress; and the notional
ingress coincides with the e2e ingress, and the notional-egress
coincides with the e2e egress. That is the case that [EL-RFC]
addresses.

Separation of control/data-plane implies that certain routers

  - Might be running software that supports signaling ELC and
    understanding an egress' ELC.
  - However, might not have the capability to insert (ELI+EL).

Such routers should not be allowed to play a spoil-sport in
preventing EL benefits for traffic traversing over them via stitched
LSPs. In other words, such routers can not act as notional-ingress or
notional-egress. However, the presence of such per-segment
ingress/egress routers on the path of a notional segment-LSP should
not prevent the notional segment-LSP from benefiting from the use of
EL.

5.2.1.1 Signaling ELC at stitching points (Translation rules)

The rules for propagating ELC, at stitching points, from a downstream
segment LSP to an upstream segment LSP are listed in this section.

There is benefit in propagating ELC across stitching points is to not
have to re-compute the EL at different segment ingress for those
segments that are EL capable, including when the LSP segments have
been setup using different protocols.

Additionally, in certain cases it should be possible to get benefits
of (ELI+EL) on LSP segments that are not "intrinsically EL capable",
where the lack of "intrinsic EL capability" is due to:
  - The per-segment ingress does not support EL insertion.
  - The per-segment PHR/egress does not support EL POPing.

However, such a stitching point might support ELC signaling.

At a stitching point, when 2 LSP segments: L1 (incoming LSP) and L2
(the outgoing LSP) are being stitched, the following rules should be

followed by stitching point in signaling its ELC.

A. Segment-egress:

   1. A segment-egress signals ELC for an LSP-segment L1 when:
      a. The segment-egress is intrinsically ELC, or
      b. When it is not intrinsically-ELC, however segment-egress for
         LSP-segment L2 (downstream of L1)- for which this
         stitching-point is segment-ingress - is signaling ELC.
         [This handles the case: Support the signaling even though it
          may not support the data-plane behavior.]

   2. A segment-egress MUST NOT signal ELC if BOTH of the following
      are true:
      a. It is also segment-ingress for another LSP-segment whose
         segment-egress is not signaling ELC.
      b. This router does not have the ability to remove an (ELI+EL)
         inserted by the segment-ingress for the LSP-segment for
         which this router is the segment-egress.

B. Segment-ingress:

   The following is relevant only for RSVP as defined in [EL-RFC].
   When this router acting as segment-egress for an LSP L1 (that is
   stitched to downstream LSP L2) is signaling ELC for L1, then this
   router must signal ELC in its Path messages using the mechanism
   defined in [EL-RFC].

   This is relevant only in the context of bidirectional LSPs.

5.2.2 Data plane aspects
5.2.2.1 Stitching: Differing EL dispositions

   At a stitching point, when 2 LSP segments: L1 (incoming LSP) and L2
   (the outgoing LSP) are being stitched, the following rules should be
   followed by the stitching point in its data-plane behavior.

   A. Rationalizing EL for the outgoing LSP segment:
      When the LSP segments L1 and L2 differ in their ELC, the stitching
      point router needs to take the following data-plane actions
      depending on its role for the e2e LSP:

        a. Notional egress behavior:
           When L1 intrinsically supports ELC and L2 does not, then
           the stitching-point router must remove the (ELI+EL), if
           present under top label, from the received data packets
           before effectively SWAPing the top label. In other words,

          in the presence of the ELI, the operations performed
          should be:

           POP(IncomingLabel), POP(ELI+EL), PUSH(OutgoingLabel)
                or alternately:
           POP, POP, SWAP(OutgoingLabel)

          Translation rule "A 2" of section 5.2.1.1 would have ensured that
           the above is doable at the stitching point.

     b. Notional ingress behavior:
        When L1 does not intrinsically support ELC and L2 does, then the
        stitching point router must POP the incoming label, insert
        (ELI+EL) before PUSHing the label for the LSP segment L2.

        The label operations performed would be:
         POP(IncomingLabel), PUSH(EL), PUSH(ELI), PUSH(OutgoingLabel),
                    or
         SWAP(EL), PUSH(ELI), PUSH(OutgoingLabel)

     c. Implicit notional ingress behavior:
        When L1 is intrinsically ELC and so is L2, the arriving data
        traffic should already have (ELI+EL) on it.

        However, it is possible that due to local configuration on the
        notional-ingress, (ELI+EL) is not being inserted. In that
        case, traffic arriving on L1 will not have (ELI+EL) on it.

        In that case, this stitching-point is the "implicit notional
        ingress" and it should insert (ELI+EL) just as if it were a
        "notional ingress".

  B. Preventing multiple (ELI+EL) pairs underneath a given forwarding
     label in the stack:

     A segment-ingress that is intrinsically-EL-capable should have
     the ability to inspect received data packets to check whether an
     (ELI+EL) already exists on the data packet underneath the top
     label.

     Not causing multiple ELs on a data packet:
          When both the LSP segments L1 and L2 support ELC, the
          stitching point router SHOULD insert an (ELI+EL) only if the
          incoming packet does not contain an (ELI+EL) underneath the
          top label. In that case, the label operations are as below:
              POP(IncomingLabel), PUSH(ELI+EL), PUSH(OutgoingLabel)

If the incoming packet already contains an (ELI+EL) underneath the
top label, an additional (ELI+EL) MUST NOT be inserted on the
packet underneath the top label that is being effectively SWAPed.

This prevents a segment ingress from inserting an (ELI+EL) when
the notional ingress has already inserted an (ELI+EL).

C. Rationalizing EL insertion (at stitching-point) for LSP hierarchy:

A stitching point router that is intrinsically-EL-capable should
have the ability to inspect received data packets to check whether
an (ELI+EL) already exists, underneath any label in the
label-stack.

If the router has such a ability, then this router MUST NOT insert
an (ELI+EL) as in "A a" above.

This helps to prevent multiple (ELI+EL)s on the packet inserted
(at a stitching point) in the context of different transport
labels in a label hierarchy.

D. Notional ingress role change at a router:

This role can change due to local configuration on the router or
due to segment egress starting/stopping to signal ELC possibly due
to a configuration change at the segment egress or due to a
configuration change at this router.
When this router becomes a notional ingress, it reacts to the
change as in "A b" above.

When this router stops being a notional ingress, this router
stops inserting the (ELI+EL) underneath the top label that this
router is
   SWAPing(if this router is stitching point), or
   PUSHING (if this router is e2e ingress).

E. Notional egress role change at a router:

This role can change due to local configuration on the router or
due the egress of a downstream stitched LSP segment starting to
signal ELC.

When this router becomes a notional egress, it reacts to the
change as in "A a" above.

When this router stops being a notional egress, this router stops

performing the label operation described in "A a" above. Instead
this router now starts to simply SWAP the top label.

5.3 New abstractions: EL applicability for LSP hierarchy
5.3.1 Management plane:

Moving the (ELI+EL) underneath a different LSP's transport label:
There are 2 ways to handle the issue of section 4.3.2:
- Configuration at the ingress LER: a configuration option should
  exist by which an operator can disable the insertion of (ELI+EL)
  on a per-LSP basis. The specific level in the LSP hierarchy for
  which to enable this configuration is based on operator knowledge
  based on:
    * Knowledge of transit routers that need EL benefits : those
      routers that have a multi-path (LAG or LSP ECMP) as egress
      interface.
    * The label hashing abilities of such routers: information
      about the specific number of labels in the label-stack that
      can be used in the hash computation; and any constraints
      about the position of the labels that can be used for
      computation when the label stack is larger than a certain
      ASIC-specific threshold.

- Configuration-based rewrite of the label stack at the ingress LER
  of an intrinsically-EL-capable LSP:

  An operator will know the forwarding characteristics (with
  regards to the number of labels that can be included in the hash
  computation) of the transit routers across the path of the e2e
  LSP that is part of an LSP hierarchy.

  By making such a configuration, the operator can ensure that the
  EL will appear in the label stack such that all transit routers
  shall be able to include the as part of the hash computation.

  The configuration would cause the label stack of the outgoing
  packet to have its extant (ELI+EL) removed, and an (ELI+EL)
  inserted just underneath the label of the LSP for which this
  ingress LER is setup to insert (ELI+EL).

5.3.2 Data plane aspects

Preventing insertion of multiple (ELI+EL)s:
At an ingress LER, the router SHOULD not insert an (ELI+EL) for an
LSP if the packet already contains an ELI.

This ensures that for a data packet on a hierarchy of LSPs, there
will be only 1 instance of an (ELI+EL). This helps to prevent the
issue of section 4.3.1.

This also ensures that when multiple LSPs in an LSP hierarchy are
intrinsically-EL-capable, the (ELI+EL) will be inserted just
underneath the transport label of the innermost LSP in the hierarchy.
However, based on section 5.3.1 there is a way by which to modify the
level in the LSP hierarchy for which an (ELI+EL) is inserted.

A more specific case of this is already covered in section "5.2.2.1
C. Rationalizing EL insertion (at stitching-point) for LSP
hierarchy:".

6. Security considerations

Security considerations as listed in section 9 of [EL-RFC] apply.

7. Acknowledgments

Many thanks to Adrian Farrel for his inputs on the stitching
scenarios, and suggesting editorial improvements.

Thanks to the EL team (Sudharsana Venkataraman, Nitin Singh, Ramji
Vijayaraghavan, Jie Yan, Abhishek Tripathi) for discussions on some
of these topics.

8. IANA considerations

None.

9 References

9.1 Normative References

[EL-RFC]   Kompella, K., Drake, J., Amante, S., Henderickx, W.,
             L. Yong, "The Use of Entropy Labels in MPLS Forwarding",
             RFC-6790, November 2012.

   [GMPLS-HIER] Kompella, K., Y. Rekhter, "Label Switched Paths (LSP)
                Hierarchy with Generalized Multi-Protocol Label
                Switching (GMPLS)", RFC-4206, October 2005.

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

   [S-MPLS]   Leymann, N., Decraene, B., Filsfils, C.,
                Konstantynowicz, M., Steinberg, D., "Seamless MPLS
                Architecture", draft-leymann-mpls-seamless-mpls,
                October 2012.

   [STITCHING] Ayyangar, A., Kompella, K., Vasseur, JP., A. Farrel,
                "Label Switched Path Stitching with Generalized
                Multiprotocol Label Switching Traffic Engineering
                (GMPLS TE)", RFC 5150, February 2008.

9.2 Informative References

   [ISSUE-DEEP] K. Kompella, "Deep Label Stacks",
                http://tools.ietf.org/agenda/84/slides/slides-84-mpls-
                15.pdf, August 2012

   [LU-BGP]   Rekhter, Y., E. Rosen, "Carrying Label Information in
                BGP-4", RFC-3107, May 2001.

Authors' Addresses

    Ravi Singh
    Juniper Networks
    1194 N. Mathilda Ave.
    Sunnyvale, CA 94089
    US

    EMail: ravis@juniper.net

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

    EMail: yshen@juniper.net

    John Drake
    Juniper Networks
    1194 N. Mathilda Ave.
    Sunnyvale, CA 94089
    US

    EMail: jdrake@juniper.net