Authors: S. Bryant   J. Drake
         University of Surrey 5GIC   Juniper Networks
         T. Li
         Juniper Networks
         Redefining ELI considered harmful; NPL considered harmful

Abstract

   Recent work on MPLS Network Actions (MNA) has produced two drafts
   that propose to redefine the MPLS Entropy Label Indicator (ELI) for
   use with MNA. [I-D.jags-mpls-ext-hdr] [I-D.gandhi-mpls-ioam] This
   work also proposes the use of a Network Programming Label (NPL) as
   another option for use with MNA. This document considers both of
   these options harmful in the sense of [GOTO].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Recent work on MPLS Network Actions (MNA) has produced two drafts
that propose to redefine the MPLS Entropy Label Indicator (ELI) for
use with MNA. [I-D.jags-mpls-ext-hdr] [I-D.gandhi-mpls-ioam] This
work also proposes the use of a Network Programming Label (NPL) as
another option for use with MNA. This document considers both of
these options harmful in the sense of [GOTO].

1.1. Requirement Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
"OPTIONAL" in this document are to be interpreted as described in
BCP14 [RFC2119] [RFC8174] when, and only when, they appear in all
capitals, as shown here.

2. On the Redefinition of ELI

In [I-D.jags-mpls-ext-hdr], there are six claims of advantages to
reusing ELI versus using a new Special Purpose Label (SPL) or a NPL.

2.1. Backward Compatiblity

An important consideration when contemplating the use of ELI is the
question of backward compatibility. There are two risks with reuse
of ELI that need to be articulated.

2.1.1. Risk of Bit Reuse

As part of the proposal to reuse ELI, the TC and TTL fields of the
Entropy Label (EL) Label Stack Entry (LSE) will be reused to provide
fields for the In-Stack Extension Header Length (IL) and Entropy
Label Control fields (ELC).

[RFC6790] says the following regarding these fields:

> The TTL for the EL MUST be zero to ensure that it is not used
> inadvertently for forwarding. The TC for the EL may be any value.

This proposal violates the first constraint. There is a small, but
not inconsequential risk that an implementation will actually check
the TTL field and change its behavior if the value is non-zero.

2.1.2. Risk of additional LSEs

[RFC6790] defines the Entropy Label as containing two LSEs, one
containing the ELI and one containing the EL.

This proposal also suggests adding additional LSEs after these two
LSEs. If a legacy Label Switch Router (LSR) sees the ELI and decides
to remove it from the label stack, then it will only remove the
first two LSEs, leaving any additional LSEs on the label stack and
effectively corrupting it, with unknown consequences.

This is a significant and unacceptable risk.

Signalling the use of the MNA label stack will avoid this problem,
but it also implies that the MNA label stack will not be seen by
legacy LSRs.

2.1.3. Risk of Bottom of Stack (BOS) data

If the MNA label stack implies that there is BOS data after the
label stack, and a legacy LSR processes the packet, then it will be
unaware of the BOS data and risks processing the BOS data as part of
the payload.

This is another significant and unacceptable risk.

2.2. Claim 1: Faster Deployment

The first claim is that reusing the ELI will speed deployment:

> Faster deployment in an existing network that has EL already
> deployed with an incremental benefit (e.g., incremental signaling
> extension for ELI capability).

To understand this claim, consider the deployment cycle for MNA. The
steps that we, as a community, must undertake are:

1. Reach rough consensus. We must determine what we will do for
   MNA. This will become a set of Internet drafts.

2. Develop software. We must write forwarding plane and signalling
   code in accordance with the above drafts. There may be some
   overlap between draft development and software development.

3. Testing. Software for a production network requires a
   significant testing effort.

4. Deployment. Even given production ready software, it must be
   deployed throughout a substantial portion of an operator
   network. To have significant field experience, multiple
   operators will have to do this in parallel. As software
   upgrades are frequently service impacting, they require
   scheduling maintenance windows and are not done frequently.
   Further, systems are typically upgraded individually, as
   failures from mass upgrades can lead to mass downtime.

Based on previous experience, this entire process is likely to take
around three to five years, depending on operator urgency. However,
the magnitude of this effort is not the issue, the claim is that
using ELI would help shorten this process.

Using ELI will not help shorten the consensus process appreciably.
There are many issues that need to be resolved. Using ELI will not
shorten the development cycle at all. Writing code for ELI or
another SPL will take the same amount of time. Similarly, there are
no advantages to reusing ELI during testing.

Finally, we must consider whether using ELI can impact the
deployment time scale. As noted in Section 2.1.2 and Section 2.1.3,
exposing an ELI label stack with added LSEs or BOS data is not
compatible with legacy LSRs. To avoid this, an operator would have
to restrict their use of the MNA label stack to only functions that
could be encoded without additional LSEs or BOS data. While this is
not impossible, it greatly limits the functionality that can be
deployed and creates an enormous operational burden on the network
operator because they must enable some, but not all MNA related

functionality. If they enable the wrong set of functions, their traffic will be lost. This seems very limiting and fragile.

This is an unacceptable combination of risk and burden.

Signaling cannot alleviate this. Signaling would direct traffic around legacy nodes, which would not be different than using a new SPL. As such, the reuse of the ELI does not seem to add significant benefits to shorten the deployment time cycle.

2.3. Claim 2: Smaller label stack

The second claim is that reusing the ELI will result in a smaller label stack:

   Single label for Entropy in the MPLS header which helps with keeping label stack size smaller.

If we use a new SPL to indicate a MNA label stack and entropy is one of the defined functions within the MNA label stack, then this is not true. There is no need to have both a MNA label stack and an ELI simultaneously. All proposals on the table are already taking this approach.

This claim is false.

2.4. Claim 3: No new hardware

The third claim is:

   When EL is already enabled in the network, the proposed scheme does not require hardware to support an additional SPL indicator.

This claim is either suggesting that (a) an additional SPL indicator is burdensome, or (b) that hardware is required to support a new SPL indicator, or (c) both.

If point (a) is the interpretation, then it must be noted that there are already many SPLs allocated and in use. One more is not a significant difference and this is a trivial claim.

If point (b) is the interpretation, then we must consider legacy hardware. Many implementations have used microcode to process the label stack. Adding an additional SPL is a microcode and not a hardware change. There are possibly some implementations that do process a label stack in pure hardware. These implementations would need a spin to support MNA, regardless of whether or not a new SPL is used. We must focus MNA deployment on the microcoded implementations.

Claim 3 is irrelevant.

2.5. Claim 4: Save an SPL

Claim 4 states:

> Save a new Special Purpose Label and related protocol extensions
> to signal its capability in LDP, RSVP-TE, BGP, IS-IS, OSPF, BGP-
> LS, etc.

This claim makes two sub-claims. First, reusing ELI would save us
from allocating another SPL. This is true, but the risks described
in Section 2.1.2 and Section 2.1.3 more than offset this small
benefit.

Second, the claim is that reusing ELI would avoid making a signaling
change. The development of MNA will already require that we make
signaling changes. To avoid backward compatibility problems, we will
end up signaling each and every function explicitly. Saving the
signaling effort for a separate SPL is inconsequential in this
light.

2.6. Claim 5: Consistent hashing

Claim 5 states:

> An intermediate node can compute ECMP hash with the EL field and
> avoid inconsistent load-balancing of traffic flow that can happen
> when MPLS Extension Header alters the label stack.

If we use a new SPL, this will also be true. The MNA substack will
contain support for an entropy action and the ISD will contain an EL
field. Transit nodes will be able to hash on this EL field.

Claim 5 is incorrect, as it not an advantage. It is exactly the same
as a new SPL.

2.7. Claim 6: Smaller label stack

Claim 6 states:

> Reduce MPLS Label stack size when EL is enabled for ECMP hashing
> when MPLS Extension Header is also used. As there is only one
> field for EL in the MPLS Header, it simplifies the MPLS header
> processing.

Assuming our MNA solution includes EL as part of its label stack,
this is always true, regardless of SPL.

Claim 6 is false.

3. On the use of Network Programming Labels (NPL)

NPL is not defined in an IETF document that the authors could find. An external reference [TDC] says:

> Network programming labels may be allocated from the global SR Global Block (SRGB) for SR Multiprotocol Label Switching (SR-MPLS) by a Software Defined Networking (SDN) controller.

This would seem to imply that an NPL is specific to an SR solution. However, the MNA requirements [I-D.bocci-mpls-miad-adi-requirements] section 3.1.1, requirement 12 says:

> Data plane mechanisms for ADIs MUST be independent of the control plane type (LDP, RSVP, BGP, static, IGP, etc).

This would seem to be contradictory.

If the NPL is an arbitrary label selected and and configured by the network operator, this would seem to be an undue burden on the operator. The purpose of standards is to avoid having unique items that must be managed intependently.

4. Conclusion

This document has shown that the use of ELI or NPL to initiate MNA processing has significant risks and limitations. While some may be willing to accept the risks on their behalf, the decisions that must be made will affect all players in the industry and must be commensurate with everyone's risk tolerance. Accordingly, reusing ELI would seem to be a poor choice and that the industry would be better served if we simply used a new SPL.

5. References

5.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[I-D.bocci-mpls-miad-adi-requirements]
                                        Bocci, M. and S. Bryant,
          "Requirements for MPLS Label Stack Indicators and
          Ancillary Data", Work in Progress, Internet-Draft, draft-
          bocci-mpls-miad-adi-requirements-02, 3 March 2022,
          <https://www.ietf.org/archive/id/draft-bocci-mpls-miad-
          adi-requirements-02.txt>.

5.2. Informative References

[I-D.jags-mpls-ext-hdr]
          Rajamanickam, J., Gandhi, R., Bhattacharya, J., Decraene,
          B., Zigler, R., Cheng, W., and L. Jalil, "MPLS Extension
          Header Encodings", Work in Progress, Internet-Draft,
          draft-jags-mpls-ext-hdr-00, 2 March 2022, <https://
          www.ietf.org/archive/id/draft-jags-mpls-ext-hdr-00.txt>.

[I-D.gandhi-mpls-ioam] Gandhi, R., Ali, Z., Brockners, F., Wen, B.,
          Decraene, B., and V. Kozak, "MPLS Data Plane
          Encapsulation for In-situ OAM Data", Work in Progress,
          Internet-Draft, draft-gandhi-mpls-ioam-04, 2 March 2022,
          <https://www.ietf.org/archive/id/draft-gandhi-mpls-
          ioam-04.txt>.

[GOTO]   Dijkstra, E. W., "Go To Statement Considered Harmful",
          DOI 10.1145/362929.362947, in Communications of The
          Association for Computing Machinery, volume 11, number 3,
          pages 147-148, 1968, <https://homepages.cwi.nl/~storm/
          teaching/reader/Dijkstra68.pdf>.

[TDC]   Filsfils, C. and R. Gandhi, "Network Programming for
          Performance and Liveness Monitoring In Segment Routing
          Networks", Technical Disclosure Commons , May 2019,
          <https://www.tdcommons.org/cgi/viewcontent.cgi?
          article=3237&context=dpubs_series>.

Authors' Addresses

Stewart Bryant
University of Surrey 5GIC

Email: sb@stewartbryant.com

John Drake
Juniper Networks

Email: jdrake@juniper.net

Tony Li
Juniper Networks

Email: tony.li@tony.li