IDR
Internet-Draft
Intended status: Standards Track
Expires: 11 November 2026
Y. Liu
S. Peng
ZTE
10 May 2026

BGP Extension for SR-MPLS Entropy Label Position
draft-ietf-idr-bgp-sr-mpls-elp-01

Abstract

This document proposes extensions for BGP to indicate the entropy label position in the SR-MPLS label stack when delivering SR Policy via BGP.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction 2
2. Conventions used in this document 3
  2.1. Requirements Language 3
  2.2. Terminology and Acronyms 3
3. Entropy Label Position in SR-MPLS with the Controller 3
4. BGP Extensions for ELP in SR Policy 5
5. Operational Considerations 7
6. IANA Considerations 8
7. Security Considerations 8
8. Acknowledgement 8
9. References 8
  9.1. Normative References 8
  9.2. Informative References 9
Authors’ Addresses 10

1. Introduction

Segment Routing (SR) leverages the source routing paradigm. Segment Routing can be instantiated on MPLS data plane which is referred to as SR-MPLS [RFC8660]. SR-MPLS leverages the MPLS label stack to construct the SR path.

Entropy labels (ELs) [RFC6790] are used in the MPLS data plane to provide entropy for load-balancing. The idea behind the entropy label is that the ingress router computes a hash based on several fields from a given packet and places the result in an additional label named "entropy label". Then, this entropy label can be used as part of the hash keys used by an LSR. Using the entropy label as part of the hash keys reduces the need for deep packet inspection in the LSR while keeping a good level of entropy in the load-balancing.

[RFC8662] proposes to use entropy labels for SR-MPLS networks and multiple <ELI, EL> pairs may be inserted in the SR-MPLS label stack. The ingress node may decide the number and position of the ELI/ELs which need to be inserted into the label stack, that is termed as ELP (Entropy Label Position) in this document. But in some cases, the controller (e.g, PCE) can be used to perform the SR path computation as well as the Entropy Label Position, which is useful for inter-domain scenarios.

[RFC9830] specifies the way to use BGP to distribute one or more of the candidate paths of an SR Policy to the headend of that policy.

This document proposes extensions for BGP to indicate the ELP in the segment list when delivering SR Policy via BGP in SR-MPLS networks.

2. Conventions used in this document

2.1. Requirements 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 BCP 14 [RFC2119][RFC8174] when, and only when, they appear in all capitals, as shown here.

2.2. Terminology and Acronyms

EL: Entropy Label

ELI: Entropy Label Indicator

ELC: Entropy Label Capability

ERLD: Entropy Readable Label Depth

ELP: Entropy Label Position

BMI: Base MPLS Imposition is the number of MPLS labels that can be imposed inclusive of all service/transport/special labels.

MSD: Maximum SID Depth

BMI-MSD: A type of MSD signals the BMI of a node.

ERLD-MSD: A type of MSD advertises the ERLD of a node.

3. Entropy Label Position in SR-MPLS with the Controller

As described in [RFC8662] section 7, ELI/EL placement is not an easy decision, multiple criteria may be taken into account.

First is the BMI-MSD[RFC8491], it defines the maximum number of labels that a particular node can impose on a packet, and it is a limit when the ingress node imposing ELI/EL pairs on the SR label stack.

The Entropy Readable Label Depth(ERLD) [RFC8662] is another important parameter to consider when inserting an ELI/EL. The ERLD is defined as the number of labels a router can both read in an MPLS packet received on its incoming interface(s) and use in its load-balancing function. An ELI/EL pair must be within the ERLD of the LSR in order for the LSR to use the EL during load-balancing. It’s necessary to get the ERLD of the nodes along the SR path to achieve efficient load-balancing.

An implementation MAY try to evaluate if load-balancing is really expected at a particular node based on the segment type of its label, which also influences the ELP of a segment list.

Other criteria includes maximizing number of LSRs that will load-balance, preference for a part of the path, and etc. Using which criteria and how to decide the ELP based on the criteria is a matter of implementation.

As shown in Figure 1, in the inter-domain scenario, a path from A to Z is required, a centralized controller performs the computation of the end-to-end path, along which traffic load-balancing is required.

```
 ....................   ....................    .....................
 .                  .   .                  .    .                   .
 .+---+       +---+ .   . +---+      +---+ .    .+---+      +----+  .
 .| A |-------| B |------ | C |------| X |-------| Y |------| Z  |  .
 .+---+       +---+ .   . +---+      +---+ .    .+---+      +----+  .
 .     domain 1     .   .      domain 2    .    .     domain 3      .
 ....................   ....................    .....................
```

Figure 1: Entropy Labels in SR-MPLS Inter-Domain Scenario

When the headend node in the first domain can’t get the information of the nodes/SIDs in other domains, e.g, the ERLD of each node or the type of the SID bounded to a node/link, it’s difficult for the headend node to decide the ELP of the segment list for the path.

Performing the computation of the ELP by the controller is an alternate, since it’s easier for the controller to get the required information along the segment list prescribed by itself.

For example, the ERLD can be advertised as ERLD-MSD via IS-IS[RFC9088] and OSPF[RFC9089] within the domain, in each domain, one or more nodes are configured with BGP-LS so the controller can get the ERLD-MSD of all the nodes through BGP-LS[RFC9088] [RFC9089]. The controller can acquire the BMI-MSD of the headend node or the Binding SID anchor node via BGP-LS[RFC8814] or PCEP[RFC8664].

Another benefit of utilizing the controller to calculate ELP is that if the criteria or calculation algorithm is changed, the corresponding modification only needs to be made on the controller instead of each headend node in the network.

When the controller performs the computation of the the ELP for a segment list, the considerations for the placement of ELI/ELs introduced in [RFC8662] are still applicable. How the controller computes the ELP is out of scope of the document.

After the ELP of an SR path is decided, the controller SHOULD inform the result to the headend node of the path, so the node knows where to insert the ELI/ELs when needed. Section 4 proposes the detailed extensions for BGP to carry this information.

4. BGP Extensions for ELP in SR Policy

As defined in [RFC9830], the SR Policy encoding structure is as follows:

```
SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
Attributes:
   Tunnel Encaps Attribute (23)
      Tunnel Type: SR Policy
          Binding SID
          Preference
          Priority
          Policy Name
          Explicit NULL Label Policy (ENLP)
          Segment List
              Weight
              Segment
              Segment
              ...
          ...
```

This document defines a new ELP sub-TLV within Segment List sub-TLV. The ELP sub-TLV has the following format:

```
 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      |   Length      |         RESERVED              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where,

Type: 17.

Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 2.

RESERVED: 2 octet of reserved bits. MUST be unset on transmission and MUST be ignored on receipt.

The ELP sub-TLV, when present, indicates that the <ELI, EL> label pair SHOULD be inserted at this position. It MAY appear multiple times in the Segment List sub-TLV.

The new SR Policy encoding structure with ELP sub-TLV is expressed as below:

```
SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
Attributes:
   Tunnel Encaps Attribute (23)
      Tunnel Type: SR Policy
          Binding SID
          Preference
          Priority
          Policy Name
          Explicit NULL Label Policy (ENLP)
          Segment List
              Weight
              Segment
              Segment
              ELP
              Segment
              ...
              Segment
              ELP
              Segment
              Segment
              ...
          ...
```

The error handling methods for the SR Policy TLVs/sub-TLVs in [RFC9830] apply to this document. The validation of each ELP sub-TLV MUST be performed in headend nodes in the BGP process to determine if they are malformed or invalid. In case of any error detected, the "treat-as-withdraw" strategy MUST be applied.

The SRPM policy determines if the values for ELP impact the selection of the SR Policy Candidate Path (SR Policy CP) as the best SR Policy CP.

When using the protocol extensions introduced in this document, scalability SHOULD be considered since it may increase the the amount of control plane information(i.e., the BGP messages) exchanged between the network controller and the headend nodes.

5. Operational Considerations

After obtaining the topolopy of the network and other necessary criterias for determining ELP, the controller calculates the path information of the SR Policy instance based on these information and builds it from the headend to the endpoint. And then the controller delivers the SR path information with the segment list as well as the EL insertion position information included to the headend, leveraging the BGP extension introduced in this document.

The <ELI, EL> label pair insertion is indicated at the segment list level. The following example shows how a headend node operates after the ELP sub-TLV is introduced.

Node A receives an SR Policy NLRI with an Segment List sub-TLV from the controller. The Segment List sub-TLV contains multiple Segment sub-TLVs and ELP sub-TLVs, e.g, <S1, S2, S3,ELP, S4, S5, S6, ELP>, it indicates that if load-balancing is required, two <ELI, EL> pairs SHOULD be inserted into the label stack of the SR-TE forwarding entry, respectively after the Label for S3 and Label for S6.

The value of EL is supplemented by the ingress node according to load-balancing function of the appropriate keys extracted from a given packet. After inserting ELI/ELs, the label stack on the ingress node would be <S1, S2, S3, ELI, EL, S4, S5, S6, ELI, EL>.

If the SR Policy is delivered without the ELP sub-TLV, the headend node MAY still insert ELI/ELs based on its own decision, as the legacy behavior of ELI/EL insertion defined in [RFC8662], making it inconvenient for network operation and troubleshooting. So it is RECOMMENDED that at at least within the same SR policy, the behavior SHOULD be consistent, that is, either the headend node inserts the ELI/ELs based on the ELP sub-TLV received or it makes its own decision regardless of the existence of the ELP sub-TLV.

The Maximum SID Depth defines the maximum number of labels that a particular node can impose on a packet. When ELI/EL insertion is required, the number of segments and ELI/ELs to be inserted in the segment list SHOULD NOT exceed the BMI-MSD limit. In this case, depending on whether ELI/EL(s) are required to be inserted and whether the headend node has sufficient BMI-MSD, the segment list delivered by the controller to the headend MAY be different. And with the additional constraints on ELI/EL insertion, the number of valid paths may be reduced, in some cases there may be no available path. Whether the ELI/EL insertion should be considered when computing the SR policy MAY be based on the configuration or local policy on the controller, or some control plane mechanism between the headend and the controller MAY be used, the details are out of the scope of this document.

6. IANA Considerations

This document defines a new sub-TLV in the registry "SR Policy List Sub-TLVs" [RFC9830] which is assigned by IANA:

Codepoint	Description	Reference
17	ELP	This document

7. Security Considerations

Security considerations in [RFC8662] and [RFC9830] apply to this document.

The ELP extension is included in the SR Policy extension [RFC9830], so it does not introduce extra security problems comparing the existing SR policy entension. SR Policies with ELP information distributed by BGP are expected to be used entirely within trusted SR domain. The ELP information is a critical piece of information about critical infrastructure. Therefore, precaution is necessary to ensure that the ELP information advertised via BGP sessions is limited to nodes in a secure manner within this trusted SR domain.

8. Acknowledgement

The authors would like to thank Ketan Talaulikar, Jie Dong and Nat Kao for their helpful review and comments. The authors would like to thank Susan Hares for her detailed shepherd review that helped in improving the document.

9. References

9.1. 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>.

[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>.

[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>.

[RFC8662] Kini, S., Kompella, K., Sivabalan, S., Litkowski, S., Shakir, R., and J. Tantsura, "Entropy Label for Source Packet Routing in Networking (SPRING) Tunnels", RFC 8662, DOI 10.17487/RFC8662, December 2019, <https://www.rfc-editor.org/info/rfc8662>.

[RFC9830] Previdi, S., Filsfils, C., Talaulikar, K., Ed., Mattes, P., and D. Jain, "Advertising Segment Routing Policies in BGP", RFC 9830, DOI 10.17487/RFC9830, September 2025, <https://www.rfc-editor.org/info/rfc9830>.

9.2. Informative References

[RFC8491] Tantsura, J., Chunduri, U., Aldrin, S., and L. Ginsberg, "Signaling Maximum SID Depth (MSD) Using IS-IS", RFC 8491, DOI 10.17487/RFC8491, November 2018, <https://www.rfc-editor.org/info/rfc8491>.

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

[RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <https://www.rfc-editor.org/info/rfc8664>.

[RFC8814] Tantsura, J., Chunduri, U., Talaulikar, K., Mirsky, G., and N. Triantafillis, "Signaling Maximum SID Depth (MSD) Using the Border Gateway Protocol - Link State", RFC 8814, DOI 10.17487/RFC8814, August 2020, <https://www.rfc-editor.org/info/rfc8814>.

[RFC9085] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Gredler, H., and M. Chen, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing", RFC 9085, DOI 10.17487/RFC9085, August 2021, <https://www.rfc-editor.org/info/rfc9085>.

[RFC9088] Xu, X., Kini, S., Psenak, P., Filsfils, C., Litkowski, S., and M. Bocci, "Signaling Entropy Label Capability and Entropy Readable Label Depth Using IS-IS", RFC 9088, DOI 10.17487/RFC9088, August 2021, <https://www.rfc-editor.org/info/rfc9088>.

[RFC9089] Xu, X., Kini, S., Psenak, P., Filsfils, C., Litkowski, S., and M. Bocci, "Signaling Entropy Label Capability and Entropy Readable Label Depth Using OSPF", RFC 9089, DOI 10.17487/RFC9089, August 2021, <https://www.rfc-editor.org/info/rfc9089>.

Authors' Addresses

Yao Liu
ZTE
Nanjing
China
Email: liu.yao71@zte.com.cn

Shaofu Peng
ZTE
Nanjing
China
Email: peng.shaofu@zte.com.cn