IDR
Internet-Draft
Intended status: Experimental
Expires: 18 September 2024

Y. Liu
S. Peng
ZTE
G. Mishra
Verizon Inc.

17 March 2024

Advertising SID Algorithm Information in BGP
draft-peng-idr-segment-routing-te-policy-attr-09

Abstract

This document defines new Segment Types and proposes extensions for BGP to provide algorithm information for SR-MPLS Adjacency-SIDs when delivering SR Policy via BGP.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Segment Routing (SR) [RFC8402] allows a headend node to steer a packet flow along any path. [RFC9256] details the concepts of SR Policy and steering into an SR Policy. These apply equally to the MPLS and IPv6 data plane instantiations of Segment Routing with their respective representations of segments as SR-MPLS SID and SRv6 SID as described in [RFC8402].

[I-D.ietf-idr-sr-policy-safi] 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. It defines a new BGP address family (SAFI), i.e., SR Policy SAFI NLRI. In UPDATE messages of that address family, the NLRI identifies an SR Policy Candidate Path, and the attributes encode the segment lists and other details of that SR Policy Candidate Path.

11 segment-descriptor types (from type A all the way to type K) for SR segments are defined [RFC9256] section 4. [I-D.ietf-idr-sr-policy-safi] specifies the encoding for segment types A and B in BGP SR Policy SAFI. And the encoding for the remaining 9 types are specified in [I-D.ietf-idr-bgp-sr-segtypes-ext].

As specified in [RFC9256], the SR algorithm can be optionally specified for Segment Types C(IPv4 Node and SID), D(IPv6 Node and SID for SR-MPLS), I(IPv6 Node and SID for SRv6), J(IPv6 Node, index for remote and local pair, and SID for SRv6), and K(IPv6 Local/Remote addresses and SID for SRv6). That is, currently the algorithm can be carried along with SR-MPLS prefix SID, SRv6 prefix SID and SRv6 adjacency SID when delivering SR Policy.

[I-D.ietf-lsr-algorithm-related-adjacency-sid] complements that, besides the SR-MPLS prefix SID, the algorithm can be also included as part of an SR-MPLS Adjacency-SID advertisement, in scenarios where multiple algorithm share the same link resource. In this case, an SR-MPLS Policy advertised to the headend may also contain algorithm specific Adjacency-SID.

This document defines new Segment Types and proposes extensions for BGP to provide algorithm information for SR-MPLS Adjacency-SIDs when delivering SR Policy via BGP.

2. 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.

3. New Segment Types for SR-MPLS Adjacency with optional SR Algorithm

This section defines four new Segment types and the corresponding Segment Sub-TLVs of Segment List Sub-TLV to provide algorithm information for SR-MPLS Adjacency-SIDs.

The processing procedures for SID with algorithm specified in [RFC9256] and [I-D.ietf-idr-bgp-sr-segtypes-ext] are still applicable for the new segment types. When the algorithm is not specified for the SID types above which optionally allow for it, the headend SHOULD use the Strict Shortest Path algorithm if available; otherwise, it SHOULD use the default Shortest Path algorithm.

3.1. Type L: IPv4 Node Address and Local Interface ID with optional SR Algorithm for SR-MPLS

This type allows for identification of an Adjacency SID or BGP Peer Adjacency SID (as defined in [RFC8402]) SR-MPLS label for point-to-point links including IP unnumbered links. The headend is required to resolve the specified IPv4 Local Node Address to the node originating it and then use the Local Interface ID to identify the point-to-point link whose adjacency is being referred to. The Local Interface ID link descriptor follows semantics as specified in [RFC9552]. This type can also be used to indicate indirection into a layer 2 interface (i.e., without IP address) like a representation of an optical transport path or a layer 2 Ethernet port or circuit at the specified node. The SR Algorithm (refer to Section 3.1.1 of [RFC8402]) MAY also be provided.

The encoding for Type L Segment Sub-TLV is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |    Length     |     Flags     | SR Algorithm  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Local Interface ID (4 octets)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 IPv4 Node Address (4 octets)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                SR-MPLS SID (optional, 4 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

Type: TBD1

Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 14 when the SR-MPLS SID is present else it MUST be 10.

Flags: 1 octet of flags as defined in [I-D.ietf-idr-bgp-sr-segtypes-ext].

SR Algorithm: 1 octet specifying SR Algorithm as described in Section 3.1.1 of [RFC8402]) when A-Flag as defined in [I-D.ietf-idr-bgp-sr-segtypes-ext] is present. SR Algorithm is used by SRPM as described in Section 4 of [RFC9256]). When A-Flag is not encoded, this field SHOULD be set to zero on transmission and MUST be ignored on receipt.

Local Interface ID: 4 octets of interface index of local interface (refer TLV 258 of [RFC9552]).

IPv4 Node Address: a 4-octet IPv4 address representing a node.

SR-MPLS SID: optional, 4-octet field containing label, TC, S and TTL as defined in Section 2.4.4.2.1 of [I-D.ietf-idr-sr-policy-safi].

3.2. Type M: IPv4 Addresses for link endpoints as Local, Remote pair with optional SR Algorithm for SR-MPLS

This type allows for identification of an Adjacency SID or BGP Peer Adjacency SID (as defined in [RFC8402]) SR-MPLS label for links. The headend is required to resolve the specified Local IPv4 Address to the node originating it and then use the Remote IPv4 Address to identify the link adjacency being referred to. The Local and Remote Address pair link descriptors follow semantics as specified in [RFC9552]. The SR Algorithm (refer to Section 3.1.1 of [RFC8402]) MAY also be provided.

The format of Type M Segment Sub-TLV is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |    Length     |     Flags     | SR Algorithm  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Local IPv4 Address (4 octets)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Remote IPv4 Address  (4 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 SR-MPLS SID (optional, 4 octets)              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

Type: TBD2

Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 14 when the SR-MPLS SID is present else it MUST be 10.

Flags: 1 octet of flags as defined in [I-D.ietf-idr-bgp-sr-segtypes-ext].

SR Algorithm: 1 octet specifying SR Algorithm as described in Section 3.1.1 of [RFC8402]) when A-Flag as defined in [I-D.ietf-idr-bgp-sr-segtypes-ext] is present. SR Algorithm is used by SRPM as described in Section 4 of [RFC9256]). When A-Flag is not encoded, this field SHOULD be set to zero on transmission and MUST be ignored on receipt.

Local IPv4 Address: a 4-octet IPv4 address representing the local link address of the node.

Remote IPv4 Address: a 4-octet IPv4 address representing the link address of the neighbor node.

SR-MPLS SID: optional, 4-octet field containing label, TC, S and TTL as defined in Section 2.4.4.2.1 of [I-D.ietf-idr-sr-policy-safi].

3.3. Type N: IPv6 Node Addresses and Interface ID for link endpoints as Local, Remote pair, with optional SR Algorithm for SR-MPLS

This type allows for identification of an Adjacency SID or BGP Peer Adjacency SID (as defined in [RFC8402]) label for links including those with only Link-Local IPv6 addresses. The headend is required to resolve the specified IPv6 Node Address to the node originating it and then use the Local Interface ID to identify the point-to-point link whose adjacency is being referred to. For other than point-to-point links, additionally the specific adjacency over the link needs to be resolved using the IPv6 Remote Node Address and Interface ID. The Local and Remote pair of Node Address and Interface ID link descriptor follows semantics as specified in [RFC9552]. This type can also be used to indicate indirection into a layer 2 interface (i.e., without IP address) like a representation of an optical transport path or a layer 2 Ethernet port or circuit at the specified node. The SR Algorithm (refer to Section 3.1.1 of [RFC8402]) MAY also be provided.

The format of Type N Segment Sub-TLV is as follows:

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |   Length      |      Flags    | SR Algorithm |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                 Local Interface ID (4 octets)                 |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    //               IPv6 Local Node Address (16 octets)           //
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                 Remote Interface ID (4 octets)                |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    //               IPv6 Remote Node Address (16 octets)          //
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                SR-MPLS SID (optional, 4 octets)               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

Type: TBD3

Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 46 when the SR-MPLS SID is present else it MUST be 42.

Flags: 1 octet of flags as defined in [I-D.ietf-idr-bgp-sr-segtypes-ext].

SR Algorithm: 1 octet specifying SR Algorithm as described in Section 3.1.1 of [RFC8402]) when A-Flag as defined in [I-D.ietf-idr-bgp-sr-segtypes-ext] is present. SR Algorithm is used by SRPM as described in Section 4 of [RFC9256]). When A-Flag is not encoded, this field SHOULD be set to zero on transmission and MUST be ignored on receipt.

Local Interface ID: 4 octets of interface index of local interface (refer TLV 258 of [RFC9552]).

IPv6 Local Node Address: a 16-octet IPv6 address representing the node.

Remote Interface ID: 4 octets of interface index of remote interface (refer TLV 258 of [RFC9552]). The value MAY be set to zero when the local node address and interface identifiers are sufficient to describe the link.

IPv6 Remote Node Address: a 16-octet IPv6 address. The value MAY be set to zero when the local node address and interface identifiers are sufficient to describe the link.

SR-MPLS SID: optional, 4-octet field containing label, TC, S and TTL as defined in Section 2.4.4.2.1 of [I-D.ietf-idr-sr-policy-safi].

3.4. Type O: IPv6 Addresses for link endpoints as Local, Remote pair, with optional SR Algorithm for SR-MPLS

This type allows for identification of an Adjacency SID or BGP Peer Adjacency SID (as defined in [RFC8402]) label for links with Global IPv6 addresses. The headend is required to resolve the specified Local IPv6 Address to the node originating it and then use the Remote IPv6 Address to identify the link adjacency being referred to. The Local and Remote IPv6 Address pair link descriptors follow semantics as specified in [RFC9552]. The SR Algorithm (refer to Section 3.1.1 of [RFC8402]) MAY also be provided.

TThe format of Type O Segment Sub-TLV is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |    Length     |     Flags     | SR Algorithm  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//               Local IPv6 Address (16 octets)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//               Remote IPv6 Address  (16 octets)              //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                SR-MPLS SID (optional, 4 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

Type: TBD4

Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 38 when the SR-MPLS SID is present else it MUST be 34.

Flags: 1 octet of flags as defined in [I-D.ietf-idr-bgp-sr-segtypes-ext].

SR Algorithm: 1 octet specifying SR Algorithm as described in Section 3.1.1 of [RFC8402]) when A-Flag as defined in [I-D.ietf-idr-bgp-sr-segtypes-ext] is present. SR Algorithm is used by SRPM as described in Section 4 of [RFC9256]). When A-Flag is not encoded, this field SHOULD be set to zero on transmission and MUST be ignored on receipt.

Local IPv6 Address: a 16-octet IPv6 address representing the local link address of the node.

Remote IPv6 Address: a 16-octet IPv6 address representing the link address of the neighbor node.

SR-MPLS SID: optional, 4-octet field containing label, TC, S and TTL as defined in Section 2.4.4.2.1 of [I-D.ietf-idr-sr-policy-safi].

4. IANA Considerations

This document requests codepoint allocations for new Sub-TLVs of the "Segment List sub-TLV" under the "BGP Tunnel Encapsulation".

Value	Description	Reference
TBD1	Segment Type L sub-TLV	This document
TBD2	Segment Type M sub-TLV	This document
TBD3	Segment Type N sub-TLV	This document
TBD4	Segment Type O sub-TLV	This document

5. Security Considerations

Procedures and protocol extensions defined in this document do not affect the security considerations discussed in [RFC9256] and [I-D.ietf-idr-sr-policy-safi].

6. Acknowledgement

The authors would like to thank Ketan Talaulikar, Nat Kao and Zhenqiang Li for their comments and suggestions.

7. References

7.1. Normative References

[I-D.ietf-idr-bgp-sr-segtypes-ext]
Talaulikar, K., Filsfils, C., Previdi, S., Mattes, P., and D. Jain, "Segment Routing Segment Types Extensions for BGP SR Policy", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-sr-segtypes-ext-03, 4 March 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-sr-segtypes-ext-03>.

[I-D.ietf-idr-sr-policy-safi]
Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., and D. Jain, "Advertising Segment Routing Policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-sr-

policy-safi-02, 16 March 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-sr-policy-safi-02>.

[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>.

[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>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

[RFC9552] Talaulikar, K., Ed., "Distribution of Link-State and Traffic Engineering Information Using BGP", RFC 9552, DOI 10.17487/RFC9552, December 2023, <https://www.rfc-editor.org/info/rfc9552>.

7.2. Informative References

[I-D.ietf-lsr-algorithm-related-adjacency-sid] Peng, S., Chen, R., Talaulikar, K., and P. Psenak, "Algorithm Related IGP-Adjacency SID Advertisement", Work in Progress, Internet-Draft, draft-ietf-lsr-algorithm-related-adjacency-sid-06, 5 December 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-lsr-algorithm-related-adjacency-sid-06>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[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>.

[RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019, <https://www.rfc-editor.org/info/rfc8665>.

[RFC8666] Psenak, P., Ed. and S. Previdi, Ed., "OSPFv3 Extensions for Segment Routing", RFC 8666, DOI 10.17487/RFC8666, December 2019, <https://www.rfc-editor.org/info/rfc8666>.

[RFC8667] Previdi, S., Ed., Ginsberg, L., Ed., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", RFC 8667, DOI 10.17487/RFC8667, December 2019, <https://www.rfc-editor.org/info/rfc8667>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

Authors' Addresses

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

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

Gyan Mishra
Verizon Inc.
Email: gyan.s.mishra@verizon.com