Network Working Group
Internet-Draft
Updates: 9012 (if approved)
Intended status: Standards Track
Expires: January 24, 2023

S. Previdi
Huawei Technologies
C. Filsfils
Cisco Systems
K. Talaulikar, Ed.
Arrcus Inc
P. Mattes
Microsoft
D. Jain
S. Lin
Google

July 23, 2022

Advertising Segment Routing Policies in BGP
draft-ietf-idr-segment-routing-te-policy-19

Abstract

This document introduces a BGP SAFI with two NLRIs to advertise a candidate path of a Segment Routing (SR) Policy. An SR Policy is an ordered list of segments (i.e., instructions) that represent a source-routed policy. An SR Policy consists of one or more candidate paths, each consisting of one or more segment lists. A headend may be provisioned with candidate paths for an SR Policy via several different mechanisms, e.g., CLI, NETCONF, PCEP, or BGP. This document specifies how BGP may be used to distribute SR Policy candidate paths. It defines sub-TLVs for the Tunnel Encapsulation Attribute for signaling information about these candidate paths.

This documents updates RFC9012 with extensions to the Color Extended Community to support additional steering modes over SR Policy.

Status of This Memo

This Internet-Draft will expire on January 24, 2023.

Copyright Notice

Table of Contents

1. Introduction 3
1.1. Requirements Language 5
2. SR Policy Encoding 6
2.1. SR Policy SAFI and NLRI 6
2.2. SR Policy and Tunnel Encapsulation Attribute 8
2.3. Remote Endpoint and Color 9
2.4. SR Policy Sub-TLVs 9
2.4.1. Preference Sub-TLV 9
2.4.2. Binding SID Sub-TLV 10
2.4.3. SRv6 Binding SID Sub-TLV 12
2.4.4. Segment List Sub-TLV 13
2.4.5. Explicit NULL Label Policy Sub-TLV 30
2.4.6. Policy Priority Sub-TLV 32
2.4.7. Policy Candidate Path Name Sub-TLV 32
2.4.8. Policy Name Sub-TLV 33
3. Color Extended Community 34
4. SR Policy Operations 36
4.1. Advertisement of SR Policies 36
4.2. Reception of an SR Policy NLRI 36
4.2.1. Acceptance of an SR Policy NLRI 36
4.2.2. Usable SR Policy NLRI 37
4.2.3. Passing a usable SR Policy NLRI to the SRPM 37
4.2.4. Propagation of an SR Policy 38
5. Error Handling and Fault Management 38
6. IANA Considerations 39
6.1. Existing Registry: Subsequent Address Family Identifiers (SAFI) Parameters 40
6.2. Existing Registry: BGP Tunnel Encapsulation Attribute Tunnel Types 40

6.3. Existing Registry: BGP Tunnel Encapsulation Attribute sub-TLVs 41
6.4. Existing Registry: Color Extended Community Flags 41
6.5. New Registry: SR Policy Segment List Sub-TLVs 41
6.6. New Registry: SR Policy Binding SID Flags 42
6.7. New Registry: SR Policy SRv6 Binding SID Flags 42
6.8. New Registry: SR Policy Segment Flags 43
6.9. New Registry: Color Extended Community Color-Only Types . 43
7. Security Considerations 44
8. Manageability Considerations 45
9. Acknowledgments 45
10. Contributors 45
11. References 46
11.1. Normative References 46
11.2. Informational References 48
Authors' Addresses 49

1. Introduction

Segment Routing (SR) [RFC8402] allows a headend node to steer a packet flow along a specific path. Intermediate per-path states are eliminated thanks to source routing.

The headend node is said to steer a flow into an SR Policy [RFC8402].

The packets steered into an SR Policy carry an ordered list of segments associated with that SR Policy.

[I-D.ietf-spring-segment-routing-policy] further details the concepts of SR Policy and steering into an SR Policy. These apply equally to the SR-MPLS and Segment Routing for IPv6 (SRv6) data-plane instantiations of Segment Routing using SR-MPLS and SRv6 Segment Identifiers (SIDs) as described in [RFC8402]. [RFC8660] describes the representation and processing of this ordered list of segments as an MPLS label stack for SR-MPLS. While [RFC8754] and [RFC8986] describe the same for SRv6 with the use of the Segment Routing Header (SRH).

The SR Policy related functionality described in [I-D.ietf-spring-segment-routing-policy] can be conceptually viewed as being incorporated in an SR Policy Module (SRPM). Following is a reminder of the high-level functionality of SRPM:

- Learning multiple candidate paths for an SR Policy via various mechanisms (CLI, NETCONF, PCEP, or BGP).

- Selection of the best candidate path for an SR Policy.

- o Associating a Binding SID (BSID) to the selected candidate path of an SR Policy.

- o Installation of the selected candidate path and its BSID in the forwarding plane.

This document specifies the use of BGP to distribute one or more of the candidate paths of an SR Policy to the headend of that policy. The document describes the functionality provided by BGP and, as appropriate, provides references for the functionality which is outside the scope of BGP (i.e. resides within SRPM on the headend node).

This document specifies a way of representing SR Policy candidate paths in BGP UPDATE messages. BGP can then be used to propagate the SR Policy candidate paths to the headend nodes in a network. The usual BGP rules for BGP propagation and best-path selection are used. At the headend of a specific policy, this will result in one or more candidate paths being installed into the "BGP table". These paths are then passed to the SRPM. The SRPM may compare them to candidate paths learned via other mechanisms and will choose one or more paths to be installed in the data plane. BGP itself does not install SR Policy candidate paths into the data plane.

This document introduces a BGP subsequent address family (SAFI) for IPv4 and IPv6 address families. In UPDATE messages of those AFI/SAFIs, the NLRI identifies an SR Policy Candidate Path while the attributes encode the segment lists and other details of that SR Policy Candidate Path.

While for simplicity we may write that BGP advertises an SR Policy, it has to be understood that BGP advertises a candidate path of an SR policy and that this SR Policy might have several other candidate paths provided via BGP (via an NLRI with a different distinguisher as defined in Section 2.1), PCEP, NETCONF, or local policy configuration.

Typically, a controller defines the set of policies and advertises them to policy headend routers (typically ingress routers). These policy advertisements use the BGP extensions defined in this document. The policy advertisement is, in most but not all cases, tailored for a specific policy headend. In this case, the advertisement may be sent on a BGP session to that headend and not propagated any further.

Alternatively, a router (i.e., a BGP egress router) advertises SR Policies representing paths to itself. In this case, it is possible

to send the policy to each headend over a BGP session to that headend, without requiring any further propagation of the policy.

An SR Policy intended only for the receiver will, in most cases, not traverse any Route Reflector (RR, [RFC4456]).

In some situations, it is undesirable for a controller or BGP egress router to have a BGP session to each policy headend. In these situations, BGP Route Reflectors may be used to propagate the advertisements. In certain other deployments, it may be necessary for the advertisement to propagate through a sequence of one or more ASes within an SR Domain (refer to Section 7 for the associated security considerations). To make this possible, an attribute needs to be attached to the advertisement that enables a BGP speaker to determine whether it is intended to be a headend for the advertised policy. This is done by attaching one or more Route Target Extended Communities to the advertisement [RFC4360].

The BGP extensions for the advertisement of SR Policies include following components:

- A Subsequent Address Family Identifier (SAFI) whose NLRIs identifies an SR Policy candidate path.

- A Tunnel Type identifier for SR Policy, and a set of sub-TLVs to be inserted into the Tunnel Encapsulation Attribute (as defined in [RFC9012]) specifying segment lists of the SR Policy candidate path, as well as other information about the SR Policy.

- One or more IPv4 address format route target extended community ([RFC4360]) attached to the SR Policy advertisement and that indicates the intended headend of such an SR Policy advertisement.

The Color Extended Community (as defined in [RFC9012]) is used to steer traffic into an SR Policy, as described in section 8.8 of [I-D.ietf-spring-segment-routing-policy]. The Section 3 of this document updates [RFC9012] with modifications to the format of the Flags field of the Color Extended Community by using the two leftmost bits of that field.

1.1. 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. SR Policy Encoding

2.1. SR Policy SAFI and NLRI

A SAFI is introduced in this document: the SR Policy SAFI with codepoint 73. The AFI used MUST be IPv4(1) or IPv6(2).

The SR Policy SAFI uses the NLRI format defined as follows:

```
+------------------+
|  NLRI Length     | 1 octet
+------------------+
|  Distinguisher   | 4 octets
+------------------+
|  Policy Color    | 4 octets
+------------------+
|  Endpoint        | 4 or 16 octets
+------------------+
```

Figure 1: SR Policy SAFI Format

where:

- o NLRI Length: 1 octet of length expressed in bits as defined in [RFC4760]. When AFI = 1 the value MUST be 96 and when AFI = 2 the value MUST be 192.

- o Distinguisher: 4-octet value uniquely identifying the policy in the context of <color, endpoint> tuple. The distinguisher has no semantic value and is solely used by the SR Policy originator to make unique (from an NLRI perspective) both for multiple candidate paths of the same SR Policy as well as candidate paths of different SR Policies (i.e. with different segment lists) with the same Color and Endpoint but meant for different headends.

- o Policy Color: 4-octet value identifying (with the endpoint) the policy. The color is used to match the color of the destination prefixes to steer traffic into the SR Policy as specified in section 8 of [I-D.ietf-spring-segment-routing-policy].

- o Endpoint: identifies the endpoint of a policy. The Endpoint may represent a single node or a set of nodes (e.g., an anycast address). The Endpoint is an IPv4 (4-octet) address or an IPv6 (16-octet) address according to the AFI of the NLRI. The address can be either a unicast or an unspecified address (0.0.0.0 for IPv4, :: for IPv6) as specified in [I-D.ietf-spring-segment-routing-policy].

The color and endpoint are used to automate the steering of BGP service routes on SR Policy as described in section 8 of[I-D.ietf-spring-segment-routing-policy].

The NLRI containing an SR Policy candidate path is carried in a BGP UPDATE message [RFC4271] using BGP multi-protocol extensions [RFC4760] with an AFI of 1 or 2 (IPv4 or IPv6) and with a SAFI of 73. The fault management and error handling in the encoding of the NLRI is specified in Section 5.

An update message that carries the MP_REACH_NLRI or MP_UNREACH_NLRI attribute with the SR Policy SAFI MUST also carry the BGP mandatory attributes. In addition, the BGP update message MAY also contain any of the BGP optional attributes.

The next-hop network address field in SR Policy SAFI (73) updates may be either a 4-octet IPv4 address or a 16-octet IPv6 address, independent of the SR Policy AFI. The length field of the next-hop address specifies the next-hop address family. If the next-hop length is 4, then the next-hop is an IPv4 address; if the next-hop length is 16, then it is a global IPv6 address; if the next-hop length is 32, then it has a global IPv6 address followed by a link-local IPv6 address. The setting of the next-hop field and its attendant processing is governed by standard BGP procedures as described in section 3 of [RFC4760] and section 3 of [RFC2545].

It is important to note that any BGP speaker receiving a BGP message with an SR Policy NLRI, will process it only if the NLRI is among the best paths as per the BGP best-path selection algorithm. In other words, this document leverages the existing BGP propagation and best-path selection rules. Details of the procedures are described in Section 4.

It has to be noted that if several candidate paths of the same SR Policy (endpoint, color) are signaled via BGP to a headend, it is RECOMMENDED that each NLRI uses a different distinguisher. If BGP has installed into the BGP table two advertisements whose respective NLRIs have the same color and endpoint, but different distinguishers, both advertisements are passed to the SRPM as different candidate paths along with their respective originator information (i.e., ASN and BGP Router-ID) as described in section 2.4 of [I-D.ietf-spring-segment-routing-policy]. The ASN would be the ASN of the origin and the BGP Router-ID is determined in the following order:

- o From the Route Origin Community [RFC4360] if present and carrying an IP Address

o As the BGP Originator ID [RFC4456] if present

o As the BGP Router-ID of the peer from which the update was received as a last resort.

2.2. SR Policy and Tunnel Encapsulation Attribute

The content of the SR Policy Candidate Path is encoded in the Tunnel Encapsulation Attribute defined in [RFC9012] using a Tunnel-Type called SR Policy Type with codepoint 15. The use of SR Policy Tunnel-type is applicable only for the AFI/SAFI pairs of (1/73, 2/73).

The SR Policy Encoding structure is as follows:

```
SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
Attributes:
   Tunnel Encapsulation Attribute (23)
      Tunnel Type: SR Policy (15)
          Binding SID
          SRv6 Binding SID
          Preference
          Priority
          Policy Name
          Policy Candidate Path Name
          Explicit NULL Label Policy (ENLP)
          Segment List
              Weight
              Segment
              Segment
              ...
          ...
```

Figure 2: SR Policy Encoding

where:

o SR Policy SAFI NLRI is defined in Section 2.1.

o Tunnel Encapsulation Attribute is defined in [RFC9012].

o Tunnel-Type is set to 15.

o Preference, Binding SID, SRv6 Binding SID, Priority, Policy Name, Policy Candidate Path Name, ENLP, Segment-List, Weight, and Segment sub-TLVs are defined in Section 2.4.

o Additional sub-TLVs may be defined in the future.

A Tunnel Encapsulation Attribute MUST NOT contain more than one TLV of type "SR Policy".

2.3. Remote Endpoint and Color

The Tunnel Egress Endpoint and Color sub-TLVs, as defined in [RFC9012], may also be present in the SR Policy encodings.

The Tunnel Egress Endpoint and Color Sub-TLVs of the Tunnel Encapsulation Attribute are not used for SR Policy encodings and therefore their value is irrelevant in the context of the SR Policy SAFI NLRI. If present, the Tunnel Egress Endpoint sub-TLV and the Color sub-TLV MUST be ignored by the BGP speaker and MUST NOT be removed from the Tunnel Encapsulation Attribute during propagation.

2.4. SR Policy Sub-TLVs

This section specifies the sub-TLVs defined for encoding the information about the SR Policy Candidate Path.

Preference, Binding SID, SRv6 Binding SID, Segment-List, Priority, Policy Name, Policy Candidate Path Name, and Explicit NULL Label Policy are the sub-TLVs introduced for the BGP Tunnel Encapsulation Attribute [RFC9012] being defined in this section.

Weight and Segment are sub-TLVs of the Segment-List sub-TLV mentioned above.

The fault management and error handling in the encoding of the sub-TLVs defined in this section are specified in Section 5.

None of the sub-TLVs defined in the following sub-sections have any effect on the BGP best-path selection or propagation procedures. These sub-TLVs are not used by the BGP path selection process and are instead passed on to SRPM as SR Policy Candidate Path information for further processing described in [I-D.ietf-spring-segment-routing-policy].

The use of SR Policy Sub-TLVs is applicable only for the AFI/SAFI pairs of (1/73, 2/73). Future documents may extend their applicability to other AFI/SAFI.

2.4.1. Preference Sub-TLV

The Preference sub-TLV is used to carry the preference of an SR Policy candidate path. The contents of this sub-TLV are used by the SRPM as described in section 2.7 of [I-D.ietf-spring-segment-routing-policy].

The Preference sub-TLV is optional and it MUST NOT appear more than once in the SR Policy encoding.

The Preference sub-TLV has 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      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Preference (4 octets)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: Preference sub-TLV

where:

o Type: 12

o Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 6.

o Flags: 1 octet of flags. None are defined in this document. Flags SHOULD be set to zero on transmission and MUST be ignored on receipt.

o RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

o Preference: a 4-octet value.

2.4.2. Binding SID Sub-TLV

The Binding SID sub-TLV is used to signal the binding SID related information of the SR Policy candidate path. The contents of this sub-TLV are used by the SRPM as described in section 6 in [I-D.ietf-spring-segment-routing-policy].

The Binding SID sub-TLV is optional and it MUST NOT appear more than once in the SR Policy encoding.

When the Binding SID sub-TLV is used to signal an SRv6 SID, the choice of its SRv6 Endpoint Behavior [RFC8986] to be instantiated is left to the headend node. It is RECOMMENDED that the SRv6 Binding SID sub-TLV defined in Section 2.4.3, that enables the specification of the SRv6 Endpoint Behavior, be used for signaling of an SRv6 Binding SID for an SR Policy candidate path.

The Binding SID 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      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Binding SID (variable, optional)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Binding SID sub-TLV

where:

o Type: 13

o Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be one of: 18 when a SRv6 BSID is present, 6 when a SR-MPLS BSID is present, or 2 when no BSID is present.

o Flags: 1 octet of flags. The following flags are defined in the registry "SR Policy Binding SID Flags" as described in Section 6.6:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|S|I|           |
+-+-+-+-+-+-+-+-+
```

Figure 5: Binding SID Flags

where:

* S-Flag: This flag encodes the "Specified-BSID-only" behavior. It is used by SRPM as described in section 6.2.3 in [I-D.ietf-spring-segment-routing-policy].

* I-Flag: This flag encodes the "Drop Upon Invalid" behavior. It is used by SRPM as described in section 8.2 in [I-D.ietf-spring-segment-routing-policy].

* The unassigned bits in the Flag octet MUST be set to zero upon transmission and MUST be ignored upon receipt.

o RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

o Binding SID: if the length is 2, then no Binding SID is present. If the length is 6 then the Binding SID is encoded in 4 octets using the format below. Traffic Class (TC), S, and TTL (Total of 12 bits) are RESERVED and MUST be set to zero and MUST be ignored.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Label                        | TC  |S|       TTL     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: Binding SID Label Encoding

If the length is 18 then the Binding SID contains a 16-octet SRv6 SID.

2.4.3. SRv6 Binding SID Sub-TLV

The SRv6 Binding SID sub-TLV is used to signal the SRv6 Binding SID related information of an SR Policy candidate path. It enables the specification of the SRv6 Endpoint Behavior [RFC8986] to be instantiated on the headend node. The contents of this sub-TLV are used by the SRPM as described in section 6 in [I-D.ietf-spring-segment-routing-policy].

The SRv6 Binding SID sub-TLV is optional. More than one SRv6 Binding SID sub-TLVs MAY be signaled in the same SR Policy encoding to indicate one or more SRv6 SIDs, each with potentially different SRv6 Endpoint Behaviors to be instantiated.

The SRv6 Binding SID 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      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 SRv6 Binding SID (16 octets)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//     SRv6 Endpoint Behavior and SID Structure (optional)     //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: SRv6 Binding SID sub-TLV

where:

o Type: 20

o Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 26 when the SRv6 Endpoint Behavior and SID Structure is present else it MUST be 18.

o Flags: 1 octet of flags. The following flags are defined in the registry "SR Policy Binding SID Flags" as described in Section 6.7:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|S|I|B|         |
+-+-+-+-+-+-+-+-+
```

Figure 8: SRv6 Binding SID Flags

where:

* S-Flag: This flag encodes the "Specified-BSID-only" behavior. It is used by SRPM as described in section 6.2.3 in [I-D.ietf-spring-segment-routing-policy].

* I-Flag: This flag encodes the "Drop Upon Invalid" behavior. It is used by SRPM as described in section 8.2 in [I-D.ietf-spring-segment-routing-policy].

* B-Flag: This flag, when set, indicates the presence of the SRv6 Endpoint Behavior and SID Structure encoding specified in Section 2.4.4.2.13.

* The unassigned bits in the Flag octet MUST be set to zero upon transmission and MUST be ignored upon receipt.

o RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

o SRv6 Binding SID: Contains a 16-octet SRv6 SID.

o SRv6 Endpoint Behavior and SID Structure: Optional, as defined in Section 2.4.4.2.13.

2.4.4. Segment List Sub-TLV

The Segment List sub-TLV encodes a single explicit path towards the endpoint as described in section 5.1 in [I-D.ietf-spring-segment-routing-policy]. The Segment List sub-TLV

includes the elements of the paths (i.e., segments) as well as an optional Weight sub-TLV.

The Segment List sub-TLV may exceed 255 bytes in length due to a large number of segments. A 2-octet length is thus required. According to section 2 of [RFC9012], the sub-TLV type defines the size of the length field. Therefore, for the Segment List sub-TLV, a code point of 128 or higher is used.

The Segment List sub-TLV is optional and MAY appear multiple times in the SR Policy encoding. The ordering of Segment List sub-TLVs, each sub-TLV encoding a Segment List, does not matter.

The Segment List sub-TLV contains zero or more Segment sub-TLVs and MAY contain a Weight sub-TLV.

The Segment List 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    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                           sub-TLVs                          //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 9: Segment List sub-TLV

where:

- o Type: 128.
- o Length: the total length (not including the Type and Length fields) of the sub-TLVs encoded within the Segment List sub-TLV in terms of octets.
- o RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.
- o sub-TLVs currently defined:
  - * An optional single Weight sub-TLV.
  - * Zero or more Segment sub-TLVs.

Validation of an explicit path encoded by the Segment List sub-TLV is beyond the scope of BGP and performed by the SRPM as described in section 5 in [I-D.ietf-spring-segment-routing-policy].

2.4.4.1. Weight Sub-TLV

The Weight sub-TLV specifies the weight associated with a given segment list. The contents of this sub-TLV are used only by the SRPM as described in section 2.11 in [I-D.ietf-spring-segment-routing-policy].

The Weight sub-TLV is optional and it MUST NOT appear more than once inside the Segment List sub-TLV.

The Weight 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      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Weight                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10: Weight sub-TLV

where:

- o Type: 9.
- o Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 8.
- o Flags: 1 octet of flags. None are defined in this document. Flags SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.
- o Weight: 4 octets value.

2.4.4.2. Segment Sub-TLVs

A Segment sub-TLV describes a single segment in a segment list (i.e., a single element of the explicit path). One or more Segment sub-TLVs constitute an explicit path of the SR Policy candidate path. The contents of these sub-TLVs are used only by the SRPM as described in section 4 in [I-D.ietf-spring-segment-routing-policy].

The Segment sub-TLVs are optional and MAY appear multiple times in the Segment List sub-TLV.

Section 4 of [I-D.ietf-spring-segment-routing-policy] defines several Segment Types:

```
Type  A: SR-MPLS Label
Type  B: SRv6 SID
Type  C: IPv4 Prefix with optional SR Algorithm
Type  D: IPv6 Global Prefix with optional SR Algorithm for SR-MPLS
Type  E: IPv4 Prefix with Local Interface ID
Type  F: IPv4 Addresses for link endpoints as Local, Remote pair
Type  G: IPv6 Prefix and Interface ID for link endpoints as Local,
         Remote pair for SR-MPLS
Type  H: IPv6 Addresses for link endpoints as Local, Remote pair
         for SR-MPLS
Type  I: IPv6 Global Prefix with optional SR Algorithm for SRv6
Type  J: IPv6 Prefix and Interface ID for link endpoints as Local,
         Remote pair for SRv6
Type  K: IPv6 Addresses for link endpoints as Local, Remote pair
         for SRv6
```

The following sub-sections specify the sub-TLVs used for encoding each of these Segment Types.

2.4.4.2.1. Segment Type A

The Type A Segment Sub-TLV encodes a single SR-MPLS SID. The format is as follows and is used to encode MPLS Label fields as specified in [RFC3032] [RFC5462].:

```
 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     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Label                        | TC  |S|       TTL     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 11: Type A Segment sub-TLV

where:

o Type: 1.

- Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 6.

- Flags: 1 octet of flags as defined in Section 2.4.4.2.12.

- RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

- Label: 20 bits of label value.

- TC: 3 bits of traffic class.

- S: 1 bit of bottom-of-stack.

- TTL: 1 octet of TTL.

The following applies to the Type-1 Segment sub-TLV:

- The S bit MUST be zero upon transmission and MUST be ignored upon reception.

- If the originator wants the receiver to choose the TC value, it sets the TC field to zero.

- If the originator wants the receiver to choose the TTL value, it sets the TTL field to 255.

- If the originator wants to recommend a value for these fields, it puts those values in the TC and/or TTL fields.

- The receiver MAY override the originator's values for these fields. This would be determined by local policy at the receiver. One possible policy would be to override the fields only if the fields have the default values specified above.

2.4.4.2.2. Segment Type B

The Type B Segment Sub-TLV encodes a single SRv6 SID. The format 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     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                       SRv6 SID (16 octets)                  //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//     SRv6 Endpoint Behavior and SID Structure (optional)     //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 12: Type B Segment sub-TLV

where:

- Type: 13.
- Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 26 when the SRv6 Endpoint Behavior and SID Structure is present else it MUST be 18.
- Flags: 1 octet of flags as defined in Section 2.4.4.2.12.
- RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.
- SRv6 SID: 16 octets of IPv6 address.
- SRv6 Endpoint Behavior and SID Structure: Optional, as defined in Section 2.4.4.2.13.

The TLV 2 defined for the advertisement of Segment Type B in the earlier versions of this document has been deprecated to avoid backward compatibility issues.

2.4.4.2.3. Segment Type C

The Type C Segment Sub-TLV encodes an IPv4 node address, SR Algorithm, and an optional SR-MPLS SID. The format 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 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 IPv4 Node Address (4 octets)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                SR-MPLS SID (optional, 4 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 13: Type C Segment sub-TLV

where:

- Type: 3.

- Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 10 when the SR-MPLS SID is present else it MUST be 6.

- Flags: 1 octet of flags as defined in Section 2.4.4.2.12.

- SR Algorithm: 1 octet specifying SR Algorithm as described in section 3.1.1 in [RFC8402] when A-Flag as defined in Section 2.4.4.2.12 is present. SR Algorithm is used by SRPM as described in section 4 in [I-D.ietf-spring-segment-routing-policy]. When A-Flag is not encoded, this field MUST be set to zero on transmission and MUST be ignored on receipt.

- 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.

2.4.4.2.4. Segment Type D

The Type D Segment Sub-TLV encodes an IPv6 node address, SR Algorithm, and an optional SR-MPLS SID. The format 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 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                 IPv6 Node Address (16 octets)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                SR-MPLS SID (optional, 4 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 14: Type D Segment sub-TLV

where:

- Type: 4
- Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 22 when the SR-MPLS SID is present else it MUST be 18.
- Flags: 1 octet of flags as defined in Section 2.4.4.2.12.
- SR Algorithm: 1 octet specifying SR Algorithm as described in section 3.1.1 in [RFC8402] when A-Flag as defined in Section 2.4.4.2.12 is present. SR Algorithm is used by SRPM as described in section 4 in [I-D.ietf-spring-segment-routing-policy]. When A-Flag is not encoded, this field MUST be set to zero on transmission and MUST be ignored on receipt.
- IPv6 Node Address: a 16-octet IPv6 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.

2.4.4.2.5. Segment Type E

The Type E Segment Sub-TLV encodes an IPv4 node address, a local interface Identifier (Local Interface ID), and an optional SR-MPLS SID. The format 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     |   RESERVED    |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                 Local Interface ID (4 octets)                 |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                 IPv4 Node Address (4 octets)                  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                SR-MPLS SID (optional, 4 octets)               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 15: Type E Segment sub-TLV

where:

o Type: 5.

o 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.

o Flags: 1 octet of flags as defined in Section 2.4.4.2.12.

o RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

o Local Interface ID: 4 octets of interface index as defined in [RFC8664].

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

o SR-MPLS SID: optional, 4-octet field containing label, TC, S and TTL as defined in Section 2.4.4.2.1.

2.4.4.2.6. Segment Type F

The Type F Segment Sub-TLV encodes an adjacency local address, an adjacency remote address, and an optional SR-MPLS SID. The format 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     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Local IPv4 Address (4 octets)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Remote IPv4 Address  (4 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                SR-MPLS SID (optional, 4 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 16: Type F Segment sub-TLV

where:

- Type: 6.
- 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 Section 2.4.4.2.12.
- RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.
- Local IPv4 Address: a 4-octet IPv4 address.
- Remote IPv4 Address: a 4-octet IPv4 address.
- SR-MPLS SID: optional, 4-octet field containing label, TC, S and TTL as defined in Section 2.4.4.2.1.

2.4.4.2.7. Segment Type G

The Type G Segment Sub-TLV encodes an IPv6 link-local adjacency with IPv6 local node address, a local interface identifier (Local Interface ID), IPv6 remote node address, a remote interface identifier (Remote Interface ID), and an optional SR-MPLS SID. The format 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     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 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)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 17: Type G Segment sub-TLV

where:

- o Type: 7
- o 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.
- o Flags: 1 octet of flags as defined in Section 2.4.4.2.12.
- o RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.
- o Local Interface ID: 4 octets of interface index as defined in [RFC8664].
- o IPv6 Local Node Address: a 16-octet IPv6 address.
- o Remote Interface ID: 4 octets of interface index as defined in [RFC8664]. The value MAY be set to zero when the local node address and interface identifiers are sufficient to describe the link.
- o 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.
- o SR-MPLS SID: optional, 4-octet field containing label, TC, S and TTL as defined in Section 2.4.4.2.1.

2.4.4.2.8. Segment Type H

The Type H Segment Sub-TLV encodes an adjacency local address, an adjacency remote address, and an optional SR-MPLS SID. The format 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     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//               Local IPv6 Address (16 octets)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//               Remote IPv6 Address  (16 octets)              //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                SR-MPLS SID (optional, 4 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 18: Type H Segment sub-TLV

where:

- o Type: 8
- o 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.
- o Flags: 1 octet of flags as defined in Section 2.4.4.2.12.
- o RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.
- o Local IPv6 Address: a 16-octet IPv6 address.
- o Remote IPv6 Address: a 16-octet IPv6 address.
- o SR-MPLS SID: optional, 4-octet field containing label, TC, S and TTL as defined in Section 2.4.4.2.1.

2.4.4.2.9. Segment Type I

The Type I Segment Sub-TLV encodes an IPv6 node address, SR Algorithm, and an optional SRv6 SID. The format 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  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                 IPv6 Node Address (16 octets)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                    SRv6 SID (optional, 16 octets)           //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//     SRv6 Endpoint Behavior and SID Structure (optional)     //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 19: Type I Segment sub-TLV

where:

o Type: 14

o Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be one of: 42 when both SRv6 SID and SRv6 Endpoint Behavior & SID Structure are present, 34 when only SRv6 SID is present, or 18 when the SRv6 SID is not present.

o Flags: 1 octet of flags as defined in Section 2.4.4.2.12.

o SR Algorithm: 1 octet specifying SR Algorithm as described in section 3.1.1 in [RFC8402] when A-Flag as defined in Section 2.4.4.2.12 is present. SR Algorithm is used by SRPM as described in section 4 in [I-D.ietf-spring-segment-routing-policy]. When A-Flag is not encoded, this field MUST be set to zero on transmission and MUST be ignored on receipt.

o IPv6 Node Address: a 16-octet IPv6 address.

o SRv6 SID: optional, a 16-octet IPv6 address.

o SRv6 Endpoint Behavior and SID Structure: Optional, as defined in Section 2.4.4.2.13.

The TLV 10 defined for the advertisement of Segment Type I in the earlier versions of this document has been deprecated to avoid backward compatibility issues.

2.4.4.2.10. Segment Type J

The Type J Segment Sub-TLV encodes an IPv6 link-local adjacency with local node address, a local interface identifier (Local Interface ID), remote IPv6 node address, a remote interface identifier (Remote Interface ID), and an optional SRv6 SID. The format 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)         //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                SRv6 SID (optional, 16 octets)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//     SRv6 Endpoint Behavior and SID Structure (optional)     //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 20: Type J Segment sub-TLV

where:

- Type: 15
- Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be one of: 66 when both SRv6 SID and SRv6 Endpoint Behavior & SID Structure are present, 58 when only SRv6 SID is present, or 42 when the SRv6 SID is not present.
- Flags: 1 octet of flags as defined in Section 2.4.4.2.12.
- SR Algorithm: 1 octet specifying SR Algorithm as described in section 3.1.1 in [RFC8402] when A-Flag as defined in Section 2.4.4.2.12 is present. SR Algorithm is used by SRPM as described in section 4 in [I-D.ietf-spring-segment-routing-policy]. When A-Flag is not encoded, this field MUST be set to zero on transmission and MUST be ignored on receipt.

- o Local Interface ID: 4 octets of interface index as defined in [RFC8664].

- o IPv6 Local Node Address: a 16-octet IPv6 address.

- o Remote Interface ID: 4 octets of interface index as defined in [RFC8664]. The value MAY be set to zero when the local node address and interface identifiers are sufficient to describe the link.

- o 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.

- o SRv6 SID: optional, a 16-octet IPv6 address.

- o SRv6 Endpoint Behavior and SID Structure: Optional, as defined in Section 2.4.4.2.13.

The TLV 11 defined for the advertisement of Segment Type J in the earlier versions of this document has been deprecated to avoid backward compatibility issues.

2.4.4.2.11. Segment Type K

The Type K Segment Sub-TLV encodes an adjacency local address, an adjacency remote address, and an optional SRv6 SID. The format 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)              //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                SRv6 SID (optional, 16 octets)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//     SRv6 Endpoint Behavior and SID Structure (optional)     //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 21: Type K Segment sub-TLV

where:

- o Type: 16

- Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be one of: 58 when both SRv6 SID and SRv6 Endpoint Behavior & SID Structure are present, 50 when only SRv6 SID is present, or 34 when the SRv6 SID is not present.

- Flags: 1 octet of flags as defined in Section 2.4.4.2.12.

- SR Algorithm: 1 octet specifying SR Algorithm as described in section 3.1.1 in [RFC8402] when A-Flag as defined in Section 2.4.4.2.12 is present. SR Algorithm is used by SRPM as described in section 4 in [I-D.ietf-spring-segment-routing-policy]. When A-Flag is not encoded, this field MUST be set to zero on transmission and MUST be ignored on receipt.

- Local IPv6 Address: a 16-octet IPv6 address.

- Remote IPv6 Address: a 16-octet IPv6 address.

- SRv6 SID: optional, a 16-octet IPv6 address.

- SRv6 Endpoint Behavior and SID Structure: Optional, as defined in Section 2.4.4.2.13.

The TLV 12 defined for the advertisement of Segment Type K in the earlier versions of this document has been deprecated to avoid backward compatibility issues.

2.4.4.2.12. Segment Flags

The Segment Types sub-TLVs described above may contain the following flags in the "Flags" field defined in Section 6.8:

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|V|A|S|B|       |
+-+-+-+-+-+-+-+-+
```

Figure 22: Segment Flags

where:

V-Flag: This flag, when set, is used by SRPM for "SID verification" as described in Section 5.1 of [I-D.ietf-spring-segment-routing-policy].

A-Flag: This flag, when set, indicates the presence of SR Algorithm id in the "SR Algorithm" field applicable to various Segment Types. SR Algorithm is used by SRPM as described in section 4 of [I-D.ietf-spring-segment-routing-policy].

S-Flag: This flag, when set, indicates the presence of the SR-MPLS or SRv6 SID depending on the segment type.

B-Flag: This flag, when set, indicates the presence of the SRv6 Endpoint Behavior and SID Structure encoding specified in Section 2.4.4.2.13.

The unassigned bits in the Flag octet MUST be set to zero upon transmission and MUST be ignored upon receipt.

The following applies to the Segment Flags:

- V-Flag applies to all Segment Types.
- A-Flag applies to Segment Types C, D, I, J, and K. If A-Flag appears with Segment Types A, B, E, F, G, and H, it MUST be ignored.
- S-Flag applies to Segment Types C, D, E, F, G, H, I, J, and K. If S-Flag appears with Segment Types A or B, it MUST be ignored.
- B-Flag applies to Segment Types B, I, J, and K. If B-Flag appears with Segment Types A, C, D, E, F, G, and H, it MUST be ignored.

2.4.4.2.13. SRv6 SID Endpoint Behavior and Structure

The Segment Types sub-TLVs described above MAY contain the SRv6 Endpoint Behavior and SID Structure [RFC8986] encoding as described below:

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Endpoint Behavior       |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    LB Length  |  LN Length    | Fun. Length   |  Arg. Length  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 23: SRv6 SID Endpoint Behavior and Structure

where:

Endpoint Behavior: 2 octets. It carries the SRv6 Endpoint Behavior code point for this SRv6 SID as defined in section 9.2 of

[RFC8986]. When set with the value 0, the choice of SRv6 Endpoint Behavior is left to the headend.

Reserved: 2 octets of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

Locator Block Length: 1 octet. SRv6 SID Locator Block length in bits.

Locator Node Length: 1 octet. SRv6 SID Locator Node length in bits.

Function Length: 1 octet. SRv6 SID Function length in bits.

Argument Length: 1 octet. SRv6 SID Arguments length in bits.

The total of the locator block, locator node, function, and argument lengths MUST be less than or equal to 128.

2.4.5. Explicit NULL Label Policy Sub-TLV

To steer an unlabeled IP packet into an SR policy, it is necessary to create a label stack for that packet, and push one or more labels onto that stack.

The Explicit NULL Label Policy (ENLP) sub-TLV is used to indicate whether an Explicit NULL Label [RFC3032] must be pushed on an unlabeled IP packet before any other labels.

If an ENLP Sub-TLV is not present, the decision of whether to push an Explicit NULL label on a given packet is a matter of local configuration.

The ENLP sub-TLV is optional and it MUST NOT appear more than once in the SR Policy encoding.

The contents of this sub-TLV are used by the SRPM as described in section 4.1 in [I-D.ietf-spring-segment-routing-policy].

```
 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     |   RESERVED    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |     ENLP      |
 +-+-+-+-+-+-+-+-+
```

Figure 24: ELNP sub-TLV

Where:

Type: 14.

Length: Specifies the length of the value field (i.e., not including Type and Length fields) in terms of octets. The value MUST be 3.

Flags: 1 octet of flags. None are defined in this document. Flags MUST be set to zero on transmission and MUST be ignored on receipt.

RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

ENLP (Explicit NULL Label Policy): Indicates whether Explicit NULL labels are to be pushed on unlabeled IP packets that are being steered into a given SR policy. This field has one of the following values:

0: Reserved.

1: Push an IPv4 Explicit NULL label on an unlabeled IPv4 packet, but do not push an IPv6 Explicit NULL label on an unlabeled IPv6 packet.

2: Push an IPv6 Explicit NULL label on an unlabeled IPv6 packet, but do not push an IPv4 Explicit NULL label on an unlabeled IPv4 packet.

3: Push an IPv4 Explicit NULL label on an unlabeled IPv4 packet, and push an IPv6 Explicit NULL label on an unlabeled IPv6 packet.

4: Do not push an Explicit NULL label.

5 - 255: Reserved.

The ENLP reserved values may be used for future extensions and implementations SHOULD ignore the ENLP Sub-TLV with these values. The behavior signaled in this Sub-TLV MAY be overridden by local configuration. The section 4.1 of [I-D.ietf-spring-segment-routing-policy] describes the behavior on the headend for the handling of the explicit null label.

2.4.6. Policy Priority Sub-TLV

An operator MAY set the Policy Priority sub-TLV to indicate the order in which the SR policies are re-computed upon topological change. The contents of this sub-TLV are used by the SRPM as described in section 2.12 in [I-D.ietf-spring-segment-routing-policy].

The Priority sub-TLV is optional and it MUST NOT appear more than once in the SR Policy encoding.

The Priority sub-TLV has 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      |  Priority     |   RESERVED    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 25: Priority sub-TLV

Where:

Type: 15

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.

Priority: a 1-octet value.

RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

2.4.7. Policy Candidate Path Name Sub-TLV

An operator MAY set the Policy Candidate Path Name sub-TLV to attach a symbolic name to the SR Policy candidate path.

Usage of Policy Candidate Path Name sub-TLV is described in section 2.6 in [I-D.ietf-spring-segment-routing-policy].

The Policy Candidate Path Name sub-TLV may exceed 255 bytes in length due to a long name. A 2-octet length is thus required. According to section 2 of [RFC9012], the sub-TLV type defines the size of the length field. Therefore, for the Policy Candidate Path Name sub-TLV a code point of 128 or higher is used.

It is RECOMMENDED that the size of the symbolic name for the candidate path is limited to 255 bytes. Implementations MAY choose to truncate long names to 255 bytes when signaling via BGP.

The Policy Candidate Path Name sub-TLV is optional and it MUST NOT appear more than once in the SR Policy encoding.

The Policy Candidate Path Name sub-TLV has 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    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 //              Policy Candidate Path Name                      //
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 26: Policy Candidate Path Name sub-TLV

Where:

Type: 129.

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

RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

Policy Candidate Path Name: Symbolic name for the SR Policy candidate path without a NULL terminator as specified in section 2.6 of [I-D.ietf-spring-segment-routing-policy].

2.4.8. Policy Name Sub-TLV

An operator MAY set the Policy Name sub-TLV to associate a symbolic name with the SR Policy for which the candidate path is being advertised via the SR Policy NLRI.

Usage of Policy Name sub-TLV is described in section 2.1 of [I-D.ietf-spring-segment-routing-policy].

The Policy Name sub-TLV may exceed 255 bytes in length due to a long policy name. A 2-octet length is thus required. According to section 2 of [RFC9012], the sub-TLV type defines the size of the length field. Therefore, for the Policy Name sub-TLV a code point of 128 or higher is used.

It is RECOMMENDED that the size of the symbolic name for the SR Policy is limited to 255 bytes. Implementations MAY choose to truncate long names to 255 bytes when signaling via BGP.

The Policy Name sub-TLV is optional and it MUST NOT appear more than once in the SR Policy encoding.

The Policy Name sub-TLV has 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    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                        Policy Name                          //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 27: Policy Name sub-TLV

Where:

Type: 130

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

RESERVED: 1 octet of reserved bits. MUST be set to zero on transmission and MUST be ignored on receipt.

Policy Name: Symbolic name for the policy. It SHOULD be a string of printable ASCII characters, without a NULL terminator.

3. Color Extended Community

The Color Extended Community [RFC9012] is used to steer traffic corresponding to BGP routes into an SR Policy with matching color value. The Color Extended Community MAY be carried in any BGP UPDATE

message whose AFI/SAFI is 1/1 (IPv4 Unicast), 2/1 (IPv6 Unicast), 1/4 (IPv4 Labeled Unicast), 2/4 (IPv6 Labeled Unicast), 1/128 (VPN-IPv4 Labeled Unicast), 2/128 (VPN-IPv6 Labeled Unicast), or 25/70 (Ethernet VPN, usually known as EVPN). Use of the Color Extended Community in BGP UPDATE messages of other AFI/SAFIs is outside the scope of this document.

Two bits from the Flags field of the Color Extended Community are used as follows to support the requirements of Color-Only steering as specified in Section 8.8 of [I-D.ietf-spring-segment-routing-policy]:

```
                     1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C O|        Unassigned         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 28: Color Extended Community Flags

The CO bits together form the Color-Only Type field which indicates the various matching criteria between BGP NH and SR Policy endpoint in addition to the matching of the color value. Following types are defined:

o Type 0: Specific Endpoint Match: Request match for the endpoint that is the BGP NH

o Type 1: Specific or Null Endpoint Match: Request match for either the endpoint that is the BGP NH or a null endpoint (e.g., like a default gateway)

o Type 2: Specific, Null, or Any Endpoint Match: Request match for either the endpoint that is the BGP NH or with a null or any endpoint

o Type 3: reserved for future use and SHOULD NOT be used. Upon reception, an implementation MUST treat it like Type 0.

The details of the SR Policy steering mechanisms based on these Color-Only types are specified in section 8.8 of [I-D.ietf-spring-segment-routing-policy].

One or more Color Extended Communities MAY be associated with a BGP route update. Sections 8.4.1, 8.5.1, and 8.8.2 of [I-D.ietf-spring-segment-routing-policy] specify the steering behaviors over SR Policies when multiple Color Extended Communities are associated with a BGP route.

4. SR Policy Operations

As mentioned in Section 1, BGP is not the actual consumer of an SR Policy NLRI. BGP is in charge of the origination and propagation of the SR Policy NLRI but its installation and use are outside the scope of BGP. The details of SR Policy installation and use are specified in [I-D.ietf-spring-segment-routing-policy].

4.1. Advertisement of SR Policies

Typically, but not limited to, an SR Policy is computed by a controller or a path computation engine (PCE) and originated by a BGP speaker on its behalf.

Multiple SR Policy NLRIs may be present with the same <color, endpoint> tuple but with different content when these SR policies are intended for different headends.

The distinguisher of each SR Policy NLRI prevents undesired BGP route selection among these SR Policy NLRIs and allows their propagation across route reflectors [RFC4456].

Moreover, one or more route targets SHOULD be attached to the advertisement, where each route target identifies one or more intended headends for the advertised SR Policy update.

If no route target is attached to the SR Policy NLRI, then it is assumed that the originator sends the SR Policy update directly (e.g., through a BGP session) to the intended receiver. In such a case, the NO_ADVERTISE community MUST be attached to the SR Policy update.

4.2. Reception of an SR Policy NLRI

On reception of an SR Policy NLRI, a BGP speaker first determines if it is acceptable and then if it is usable.

4.2.1. Acceptance of an SR Policy NLRI

When a BGP speaker receives an SR Policy NLRI from a neighbor it MUST first, determine if it is acceptable. The following rules apply in addition to the validation described in Section 5:

- o The SR Policy NLRI MUST include a distinguisher, color, and endpoint field which implies that the length of the NLRI MUST be either 12 or 24 octets (depending on the address family of the endpoint).

- o The SR Policy update MUST have either the NO_ADVERTISE community or at least one route target extended community in IPv4-address format or both. If a router supporting this specification receives an SR Policy update with no route target extended communities and no NO_ADVERTISE community, the update MUST be considered as malformed.

- o The Tunnel Encapsulation Attribute MUST be attached to the BGP Update and MUST have a Tunnel Type TLV set to SR Policy (codepoint is 15).

A router that receives an SR Policy update that is not valid according to these criteria MUST treat the update as malformed and the SR Policy candidate path MUST NOT be passed to the SRPM.

4.2.2. Usable SR Policy NLRI

An SR Policy update that has been determined to be acceptable is further evaluated for its usability by the receiving node.

An SR Policy NLRI update without any route target extended community but having the NO_ADVERTISE community is considered usable.

If one or more route targets are present, then at least one route target MUST match the BGP Identifier of the receiver for the update to be considered usable. The BGP Identifier is defined in [RFC4271] as a 4-octet IPv4 address. Therefore, the route target extended community MUST be of the same format.

If one or more route targets are present and none matches the local BGP Identifier, then, while the SR Policy NLRI is acceptable, it is not usable on the receiver node.

When the SR Policy tunnel type includes any sub-TLV that is unrecognized or unsupported, the update SHOULD NOT be considered usable. An implementation MAY provide an option for ignoring unsupported sub-TLVs.

4.2.3. Passing a usable SR Policy NLRI to the SRPM

Once BGP on the receiving node has determined that the SR Policy NLRI is usable, it passes the SR Policy candidate path to the SRPM. Note that, along with the candidate path details, BGP also passes the originator information for breaking ties in the candidate path selection process as described in section 2.4 in [I-D.ietf-spring-segment-routing-policy].

When an update for an SR Policy NLRI results in its becoming unusable, BGP MUST delete its corresponding SR Policy candidate path from the SRPM.

The SRPM applies the rules defined in section 2 in [I-D.ietf-spring-segment-routing-policy] to determine whether the SR Policy candidate path is valid and to select the best candidate path among the valid ones for a given SR Policy.

4.2.4. Propagation of an SR Policy

SR Policy NLRIs that have been determined acceptable and valid can be evaluated for propagation, even the ones that are not usable.

SR Policy NLRIs that have the NO_ADVERTISE community attached to them MUST NOT be propagated.

By default, a BGP node receiving an SR Policy NLRI MUST NOT propagate it to any EBGP neighbor. An implementation MAY provide an explicit configuration to override this and enable the propagation of acceptable SR Policy NLRIs to specific EBGP neighbors.

A BGP node advertises a received SR Policy NLRI to its IBGP neighbors according to normal IBGP propagation rules.

By default, a BGP node receiving an SR Policy NLRI SHOULD NOT remove route target extended community before propagation. An implementation MAY provide support for configuration to filter and/or remove route target extended community before propagation.

A BGP node MUST NOT alter the SR Policy information carried in the Tunnel Encapsulation Attribute during propagation.

5. Error Handling and Fault Management

This section describes the error handling actions, as described in [RFC7606], that are to be performed for the handling of the BGP update messages for BGP SR Policy SAFI.

A BGP Speaker MUST perform the following syntactic validation of the SR Policy NLRI to determine if it is malformed. This includes the validation of the length of each NLRI and the total length of the MP_REACH_NLRI and MP_UNREACH_NLRI attributes. It also includes the validation of the consistency of the NLRI length with the AFI and the endpoint address as specified in Section 2.1.

When the error determined allows for the router to skip the malformed NLRI(s) and continue the processing of the rest of the update

message, then it MUST handle such malformed NLRIs as 'Treat-as-withdraw'. In other cases, where the error in the NLRI encoding results in the inability to process the BGP update message (e.g. length related encoding errors), then the router SHOULD handle such malformed NLRIs as 'AFI/SAFI disable' when other AFI/SAFI besides SR Policy are being advertised over the same session. Alternately, the router MUST perform 'session reset' when the session is only being used for SR Policy or when it 'AFI/SAFI disable' action is not possible.

The validation of the TLVs/sub-TLVs introduced in this document and defined in their respective sub-sections of Section 2.4 MUST be performed to determine if they are malformed or invalid. The validation of the Tunnel Encapsulation Attribute itself and the other TLVs/sub-TLVs specified in [RFC9012] MUST be done as described in that document. In case of any error detected, either at the attribute or its TLV/sub-TLV level, the "treat-as-withdraw" strategy MUST be applied. This is because an SR Policy update without a valid Tunnel Encapsulation Attribute (comprising of all valid TLVs/sub-TLVs) is not usable.

An SR Policy update that is determined to be not acceptable, and therefore malformed, based on rules described in Section 4.2.1 MUST be handled by the "treat-as-withdraw" strategy.

The validation of the individual fields of the TLVs/sub-TLVs defined in Section 2.4 are beyond the scope of BGP as they are handled by the SRPM as described in the individual TLV/sub-TLV sub-sections. A BGP implementation MUST NOT perform semantic verification of such fields nor consider the SR Policy update to be invalid or not acceptable/usable based on such validation.

An implementation SHOULD log any errors found during the above validation for further analysis.

6. IANA Considerations

This document uses code point allocations from the following existing registries:

- Subsequent Address Family Identifiers (SAFI) Parameters registry

- BGP Tunnel Encapsulation Attribute Tunnel Types registry under the BGP Tunnel Encapsulation registry

- BGP Tunnel Encapsulation Attribute sub-TLVs registry under the BGP Tunnel Encapsulation registry

- o Color Extended Community Flags registry under the BGP Tunnel Encapsulation registry

This document also requests the creation of the following new registries:

- o SR Policy Segment List Sub-TLVs under the BGP Tunnel Encapsulation registry

- o SR Policy Binding SID Flags under the BGP Tunnel Encapsulation registry

- o SR Policy SRv6 Binding SID Flags under the BGP Tunnel Encapsulation registry

- o SR Policy Segment Flags under the BGP Tunnel Encapsulation registry

- o Color Extended Community Color-Only Types registry under the BGP Tunnel Encapsulation registry

6.1. Existing Registry: Subsequent Address Family Identifiers (SAFI) Parameters

This document introduces a SAFI in the registry "Subsequent Address Family Identifiers (SAFI) Parameters" that has been assigned a code point by IANA as follows:

Code Point	Description	Reference
73	SR Policy SAFI	This document

Table 1: BGP SAFI Code Point

6.2. Existing Registry: BGP Tunnel Encapsulation Attribute Tunnel Types

This document introduces a Tunnel-Type in the registry "BGP Tunnel Encapsulation Attribute Tunnel Types" that has been assigned a codepoint by IANA as follows:

Code Point	Description	Reference
15	SR Policy	This document

Table 2: Tunnel Type Code Point

6.3. Existing Registry: BGP Tunnel Encapsulation Attribute sub-TLVs

This document defines sub-TLVs in the registry "BGP Tunnel Encapsulation Attribute sub-TLVs" that have been assigned code points by IANA as follows via the early allocation process which needs to be made permanent:

Code Point	Description	Reference
12	Preference sub-TLV	This document
13	Binding SID sub-TLV	This document
14	ENLP sub-TLV	This document
15	Priority sub-TLV	This document
20	SRv6 Binding SID sub-TLV	This document
128	Segment List sub-TLV	This document
129	Policy Candidate Path Name sub-TLV	This document
130	Policy Name sub-TLV	This document

Table 3: BGP Tunnel Encapsulation Attribute Code Points

6.4. Existing Registry: Color Extended Community Flags

This document defines the use of 2 bits in the registry called "Color Extended Community Flags" under the "BGP Tunnel Encapsulation" registry that have been assigned by IANA via the early allocation process to form the Color-Only Types field which needs to be made permanent:

Bit Position	Description	Reference
0-1	Color-only Types Field	This document

Table 4: Color Extended Community Flag Bits

6.5. New Registry: SR Policy Segment List Sub-TLVs

This document requests the creation of a new registry called "SR Policy Segment List Sub-TLVs" under the "BGP Tunnel Encapsulation" registry. The allocation policy of this registry is "Standards Action" according to [RFC8126].

Following initial Sub-TLV codepoints are assigned by this document:

Value	Description	Reference
0	Reserved	This document
1	Segment Type A sub-TLV	This document
2	Deprecated	This document
3	Segment Type C sub-TLV	This document
4	Segment Type D sub-TLV	This document
5	Segment Type E sub-TLV	This document
6	Segment Type F sub-TLV	This document
7	Segment Type G sub-TLV	This document
8	Segment Type H sub-TLV	This document
9	Weight sub-TLV	This document
10	Deprecated	This document
11	Deprecated	This document
12	Deprecated	This document
13	Segment Type B sub-TLV	This document
14	Segment Type I sub-TLV	This document
15	Segment Type J sub-TLV	This document
16	Segment Type K sub-TLV	This document
17-255	Unassigned	

Table 5: SR Policy Segment List Code Points

6.6. New Registry: SR Policy Binding SID Flags

This document requests the creation of a new registry called "SR Policy Binding SID Flags" under the "BGP Tunnel Encapsulation" registry. The allocation policy of this registry is "Standards Action" according to [RFC8126].

The following flags are defined:

Bit	Description	Reference
0	Specified-BSID-Only Flag (S-Flag)	This document
1	Drop Upon Invalid Flag (I-Flag)	This document
2-7	Unassigned	

Table 6: SR Policy Binding SID Flags

6.7. New Registry: SR Policy SRv6 Binding SID Flags

This document requests the creation of a new registry called "SR Policy SRv6 Binding SID Flags" under the "BGP Tunnel Encapsulation" registry. The allocation policy of this registry is "Standards Action" according to [RFC8126].

The following flags are defined:

Bit	Description	Reference
0	Specified-BSID-Only Flag (S-Flag)	This document
1	Drop Upon Invalid Flag (I-Flag)	This document
2	SRv6 Endpoint Behavior & SID Structure Flag (B-Flag)	This document
3-7	Unassigned	

Table 7: SR Policy SRv6 Binding SID Flags

6.8. New Registry: SR Policy Segment Flags

This document requests the creation of a new registry called "SR Policy Segment Flags" under the "BGP Tunnel Encapsulation" registry. The allocation policy of this registry is "Standards Action" according to [RFC8126].

The following flags are defined:

Bit	Description	Reference
0	Segment Verification Flag (V-Flag)	This document
1	SR Algorithm Flag (A-Flag)	This document
2	SID Specified Flag (S-Flag)	This document
3	SRv6 Endpoint Behavior & SID Structure Flag (B-Flag)	This document
4-7	Unassigned	

Table 8: SR Policy Segment Flags

6.9. New Registry: Color Extended Community Color-Only Types

This document requests the creation of a new registry called "Color Extended Community Color-Only Types" under the "BGP Tunnel Encapsulation" registry for assignment of codepoints (values 0 through 3) in the Color-Only Type field of the Color Extended Community Flags field. The allocation policy of this registry is "Standards Action" according to [RFC8126].

The following types are defined:

Type	Description	Reference
0	Specific Endpoint Match	This document
1	Specific or Null Endpoint Match	This document
2	Specific, Null, or Any Endpoint Match	This document
3	Unassigned	This document

Table 9: Color Extended Community Color-Only Types

7. Security Considerations

The security mechanisms of the base BGP security model apply to the extensions described in this document as well. See the Security Considerations section of [RFC4271] for a discussion of BGP security. Also, refer to [RFC4272] and [RFC6952] for analysis of security issues for BGP.

The BGP SR Policy extensions specified in this document enable traffic engineering and service programming use-cases within an SR domain as described in [I-D.ietf-spring-segment-routing-policy]. SR operates within a trusted SR domain [RFC8402] and its security considerations also apply to BGP sessions when carrying SR Policy information. The SR Policies distributed by BGP are expected to be used entirely within this trusted SR domain, i.e., within a single AS or between multiple ASes/domains within a single provider network. Therefore, precaution is necessary to ensure that the SR Policy information advertised via BGP sessions is limited to nodes in a secure manner within this trusted SR domain. BGP peering sessions for address-families other than SR Policy SAFI may be set up to routers outside the SR domain. The isolation of BGP SR Policy SAFI peering sessions may be used to ensure that the SR Policy information is not advertised by accident or error to an EBGP peering session outside the SR domain.

Additionally, it may be considered that the export of SR Policy information, as described in this document, constitutes a risk to confidentiality of mission-critical or commercially sensitive information about the network (more specifically endpoint/node addresses, SR SIDs, and the SR Policies deployed). BGP peerings are not automatic and require configuration; thus, it is the responsibility of the network operator to ensure that only trusted nodes (that include both routers and controller applications) within the SR domain are configured to receive such information.

8. Manageability Considerations

The specification of BGP models is an ongoing work based on [I-D.ietf-idr-bgp-model] and its future extensions are expected to cover the SR Policy SAFI. Existing BGP operational procedures also apply to the SAFI specified in this document. The management, operations, and monitoring of BGP speakers and the SR Policy SAFI sessions between them are not very different from other BGP sessions and can be managed using the same data models.

The YANG model for the operation and management of SR Policies [I-D.ietf-spring-sr-policy-yang] reports the SR Policies provisioned via BGP SR Policy SAFI along with their operational states.

9. Acknowledgments

The authors of this document would like to thank Shyam Sethuram, John Scudder, Przemyslaw Krol, Alex Bogdanov, Nandan Saha, Bruno Decraene, Gurusiddesh Nidasesi, Kausik Majumdar, Zafar Ali, Swadesh Agarwal, Jakob Heitz, Viral Patel, Peng Shaofu, Cheng Li, Martin Vigoureux, John Scudder, Vincent Roca, Brian Haberman, and Mohamed Boucadair for their comments and review of this document. The authors would like to thank Sue Hares for her detailed shepherd review that helped in improving the document.

10. Contributors

Eric Rosen
Juniper Networks
US

Email: erosen@juniper.net

Arjun Sreekantiah
Cisco Systems
US

Email: asreekan@cisco.com

Acee Lindem
Cisco Systems
US

Email: acee@cisco.com

Siva Sivabalan
Cisco Systems
US

Email: msiva@cisco.com

Imtiyaz Mohammad
Arista Networks
India

Email: imtiyaz@arista.com

Gaurav Dawra
Cisco Systems
US

Email: gdawra.ietf@gmail.com

Peng Shaofu
ZTE Corporation
China

Email: peng.shaofu@zte.com.cn

11. References

11.1. Normative References

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-22 (work in progress), March 2022.

[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>.

[RFC2545] Marques, P. and F. Dupont, "Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing", RFC 2545, DOI 10.17487/RFC2545, March 1999, <https://www.rfc-editor.org/info/rfc2545>.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <https://www.rfc-editor.org/info/rfc4360>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[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>.

[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>.

[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>.

[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>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

[RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <https://www.rfc-editor.org/info/rfc9012>.

11.2. Informational References

[I-D.ietf-idr-bgp-model]
Jethanandani, M., Patel, K., Hares, S., and J. Haas, "BGP YANG Model for Service Provider Networks", draft-ietf-idr-bgp-model-14 (work in progress), July 2022.

[I-D.ietf-spring-sr-policy-yang]
Raza, K., Sawaya, R., Shunwan, Z., Voyer, D., Durrani, M., Matsushima, S., and V. P. Beeram, "YANG Data Model for Segment Routing Policy", draft-ietf-spring-sr-policy-yang-01 (work in progress), April 2021.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <https://www.rfc-editor.org/info/rfc4272>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013, <https://www.rfc-editor.org/info/rfc6952>.

Authors' Addresses

Stefano Previdi
Huawei Technologies
IT

Email: stefano@previdi.net

Clarence Filsfils
Cisco Systems
Brussels
BE

Email: cfilsfil@cisco.com

Ketan Talaulikar (editor)
Arrcus Inc
India

Email: ketant.ietf@gmail.com

Paul Mattes
Microsoft
One Microsoft Way
Redmond, WA 98052
USA

Email: pamattes@microsoft.com

Dhanendra Jain
Google

Email: dhanendra.ietf@gmail.com

Steven Lin
Google

Email: stevenlin@google.com