

Network Working Group S.
Previdi
Internet-Draft
Individual
Intended status: Standards Track C.
Filsfils
Expires: May 17, 2021 K. Talaulikar,
Ed. Cisco
Systems P.
Mattes
Microsoft E.
Rosen Juniper
Networks D.
Jain S.
Lin
Google
2020 November 13,

**Advertising Segment Routing Policies in BGP
draft-ietf-idr-segment-routing-te-policy-11**

Abstract

This document defines a new BGP SAFI with a new NLRI in order to advertise a candidate path of a Segment Routing (SR) Policy. An SR Policy is a set of candidate paths, each consisting of one or more segment lists. The headend of an SR Policy may learn multiple candidate paths for an SR Policy. Candidate paths may be learned via a number of different mechanisms, e.g., CLI, NetConf, PCEP, or BGP. This document specifies the way in which BGP may be used to distribute SR Policy candidate paths. New sub-TLVs for the Tunnel Encapsulation Attribute are defined for signaling information about these candidate paths.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six

months

and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 17, 2021.

Previdi, et al.
1]

Expires May 17, 2021

[Page

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

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	7
2.3.	Remote Endpoint and Color	8
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	11
2.4.4.	Segment List Sub-TLV	13
2.4.5.	Explicit NULL Label Policy Sub-TLV	28
2.4.6.	Policy Priority Sub-TLV	29
2.4.7.	Policy Candidate Path Name Sub-TLV	30
2.4.8.	Policy Name Sub-TLV	31
3.	Color Extended Community	31
4.	SR Policy Operations	32

32	4.1.	Advertisement of SR Policies
32	4.2.	Reception of an SR Policy NLRI
33	4.2.1.	Acceptance of an SR Policy NLRI
33	4.2.2.	Usable SR Policy NLRI
34	4.2.3.	Passing a usable SR Policy NLRI to the SRPM
34	4.2.4.	Propagation of an SR Policy
34	5.	Error Handling
35	6.	IANA Considerations
36	6.1.	Existing Registry: Subsequent Address Family Identifiers (SAFI) Parameters
36	6.2.	Existing Registry: BGP Tunnel Encapsulation Attribute Tunnel Types
36	6.3.	Existing Registry: BGP Tunnel Encapsulation Attribute sub-TLVs

- [6.4.](#) Existing Registry: Color Extended Community Flags
[37](#)
- [6.5.](#) New Registry: SR Policy Segment List Sub-TLVs
[37](#)
- [6.6.](#) New Registry: SR Policy Binding SID Flags
[38](#)
- [6.7.](#) New Registry: SR Policy SRv6 Binding SID Flags
[38](#)
- [6.8.](#) New Registry: SR Policy Segment Flags
[39](#)
- [6.9.](#) Guidance for Designated Experts
[39](#)
- [7.](#) Security Considerations
[39](#)
- [8.](#) Acknowledgments
[40](#)
- [9.](#) Contributors
[40](#)
- [10.](#) References
[41](#)
- [10.1.](#) Normative References
[41](#)
- [10.2.](#) Informational References
[42](#)
- [Appendix A.](#) Deprecated Segment Sub-TLVs
[43](#)
- [A.1.](#) Type B-Deprecated: SID only, in the form of IPv6 address
[43](#)
- [A.2.](#) Type I-Deprecated: IPv6 Node Address with optional SRv6
SID
[43](#)
- [A.3.](#) Type J-Deprecated: IPv6 Address + Interface ID for local
and remote pair for SRv6 with optional SID
[44](#)
- [A.4.](#) Type K-Deprecated: IPv6 Local and Remote addresses for
SRv6 with optional SID
[46](#)
- [Authors' Addresses](#)
[46](#)

[1.](#) Introduction

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

The headend node is said to steer a flow into a SR Policy.

The header of a packet steered in an SR Policy is augmented with the ordered list of segments associated with that SR Policy.

[I-D.ietf-spring-segment-routing-policy] details the concepts of SR

Policy and steering into an SR Policy. These apply equally to the MPLS and IPv6 (known as SRv6) data plane instantiations of Segment Routing with their respective representations of segments as SR-MPLS SID and SRv6 SID as described in [[RFC8402](#)].

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:

- o Learning multiple candidate paths for an SR Policy via various mechanisms (CLI, NetConf, PCEP or BGP).
- o Selection of the best candidate path for an SR Policy.

- o Binding 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 way to use 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 that 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 the 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 defines a new BGP address family (SAFI). 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.

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 this document), PCEP, NETCONF or local policy configuration.

Typically, a controller defines the set of policies and advertise them to policy head-end routers (typically ingress routers). The policy advertisement uses BGP extensions defined in this document. The policy advertisement is, in most but not all of the cases, tailored for a specific policy head-end. In this case the advertisement may be sent on a BGP session to that head-end 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 head-end over a BGP session to that head-end, without requiring any further propagation of the policy.

Previdi, et al.
4]

Expires May 17, 2021

[Page

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 head-end. In these situations, BGP Route Reflectors may be used to propagate the advertisements, or it may be necessary for the advertisement to propagate through a sequence of one or more AS. 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 head-end 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:

- o A new Subsequent Address Family Identifier (SAFI) whose NLRI identifies an SR Policy candidate path.
- o A new Tunnel Type identifier for SR Policy, and a set of sub-TLVs to be inserted into the Tunnel Encapsulation Attribute (as defined in [[I-D.ietf-idr-tunnel-encaps](#)]) specifying segment lists of the SR Policy candidate path, as well as other information about the SR Policy.
- o One or more IPv4 address format route-target extended community ([RFC4360](#)) attached to the SR Policy advertisement and that indicates the intended head-end of such SR Policy advertisement.
- o The Color Extended Community (as defined in [[I-D.ietf-idr-tunnel-encaps](#)]) and used in order to steer traffic into an SR Policy, as described in section 8.4 in [[I-D.ietf-spring-segment-routing-policy](#)]. This document ([Section 3](#)) modifies the format of the Color Extended Community by using the two leftmost bits of the RESERVED 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 new SAFI is defined: the SR Policy SAFI with codepoint 73. The AFI used MUST be IPv4(1) or IPv6(2).

The SR Policy SAFI uses a new NLRI defined as follows:

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

where:

- o NLRI Length: 1 octet of length expressed in bits as defined in [RFC4760]. When AFI = 1 value MUST be 96 and when AFI = 2 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) multiple candidate paths of the same SR Policy.
- 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 [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 color and endpoint are used to automate the steering of BGP Payload prefixes on SR Policy as described in [I-D.ietf-spring-segment-routing-policy].

The NLRI containing the 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.

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; and 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 in \[RFC4760\]](#).

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 head-end, it is RECOMMENDED that each NLRI use 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 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 [[I-D.ietf-idr-tunnel-encaps](#)]
using
a new Tunnel-Type called SR Policy Type with codepoint 15.

Previdi, et al.
7]

Expires May 17, 2021

[Page

The SR Policy Encoding structure is as follows:

SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>

Attributes:

```
Tunnel Encaps Attribute (23)
  Tunnel Type: SR Policy
    Binding SID
    SRv6 Binding SID
    Preference
    Priority
    Policy Name
    Policy Candidate Path Name
    Explicit NULL Label Policy (ENLP)
    Segment List
      Weight
      Segment
      Segment
      ...
  ...
```

where:

- o SR Policy SAFI NLRI is defined in [Section 2.1](#).
- o Tunnel Encapsulation Attribute is defined in [\[I-D.ietf-idr-tunnel-encaps\]](#).
- 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 this document.
- 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 Remote Endpoint and Color sub-TLVs, as defined in [\[I-D.ietf-idr-tunnel-encaps\]](#), MAY also be present in the SR Policy encodings.

The Remote 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 Remote Endpoint sub-TLV and the Color sub-TLV MUST be ignored by the BGP speaker.

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 new sub-TLVs of the BGP Tunnel Encapsulation Attribute

[[I-D.ietf-idr-tunnel-encaps](#)] being defined in this section.

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

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 BGP 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](#)] .

2.4.1. Preference Sub-TLV

The Preference sub-TLV is used to carry the preference of the SR Policy candidate path. The contents of this sub-TLV are used by the SRPM as described in section 2.7 in [[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	2	3	4	5	6	7	8	9								
Type												Length						Flags						RESERVED																							
Preference (4 octets)																																															

where:

- o Type: 12
- o Length: 6.
- o Flags: 1 octet of flags. None are defined at this stage. Flags SHOULD be set to zero on transmission and MUST be ignored on receipt.

- o RESERVED: 1 octet of reserved bits. SHOULD 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

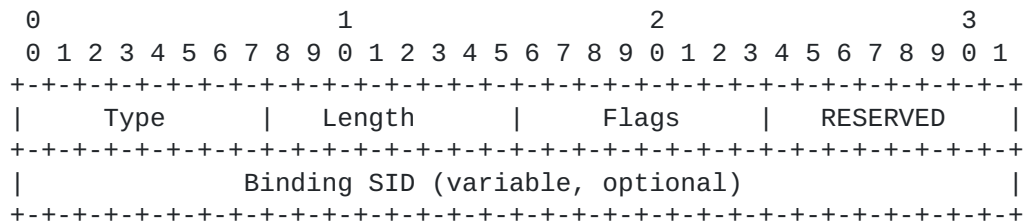
The Binding SID sub-TLV is used to signal the binding SID related information of the SR Policy candidate path. The contents of this sub-TLV are used by the SRPM as described in section 6 in [\[I-D.ietf-spring-segment-routing-policy\]](#).

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 [\[I-D.ietf-spring-srv6-network-programming\]](#) to be instantiated is left

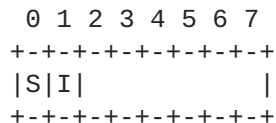
to the headend node. It is RECOMMENDED that the SRv6 Binding SID sub-TLV defined in [Section 2.4.3](#), that enables the specification of the SRv6 Endpoint Behavior, be used for signaling of an SRv6 Binding SID for an SR Policy candidate path.

The Binding SID sub-TLV has the following format:



where:

- o Type: 13
- o Length: specifies the length of the value field not including Type and Length fields. Can be 2 or 6 or 18.
- o Flags: 1 octet of flags. Following flags are defined in the new registry "SR Policy Binding SID Flags" as described in [Section 6.6](#):



where:

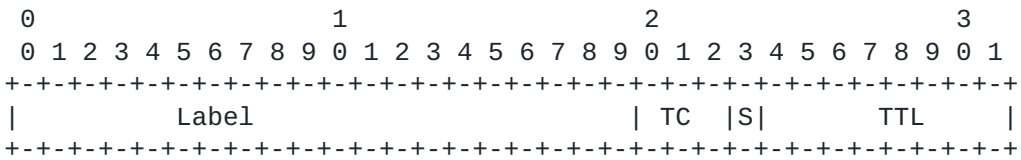
* S-Flag: This flag encodes the "Specified-BSID-only" behavior. It is used by SRPM as described in section 6.2.3 in [\[I-D.ietf-spring-segment-routing-policy\]](#).

It

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

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

- o RESERVED: 1 octet of reserved bits. SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o Binding SID: if length is 2, then no Binding SID is present. If length is 6 then the Binding SID is encoded in 4 octets using the format below. TC, S, TTL (Total of 12 bits) are RESERVED and SHOULD be set to zero and MUST be ignored.



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

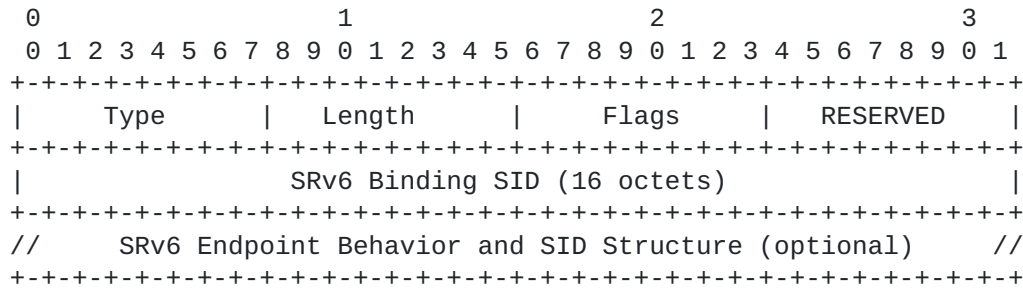
2.4.3. SRv6 Binding SID Sub-TLV

The SRv6 Binding SID sub-TLV is used to signal the SRv6 Binding SID related information of the SR Policy candidate path. It enables the specification of the SRv6 Endpoint Behavior [\[I-D.ietf-spring-srv6-network-programming\]](#) 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

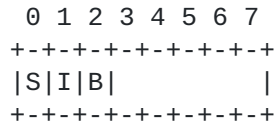
SIDs MAY be signalled in the same SR Policy encoding to indicate one or more SRv6 SIDs, each with potentially different SRv6 Endpoint Behaviors to be instantiated.

The SRv6 Binding SID sub-TLV has the following format:



where:

- o Type: TBD
- o Length is variable
- o Flags: 1 octet of flags. Following flags are defined in the new registry "SR Policy Binding SID Flags" as described in [Section 6.7](#):



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](#).
- * Unused bits in the Flag octet SHOULD be set to zero upon transmission and MUST be ignored upon receipt.

It

SRv6

- o RESERVED: 1 octet of reserved bits. SHOULD 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

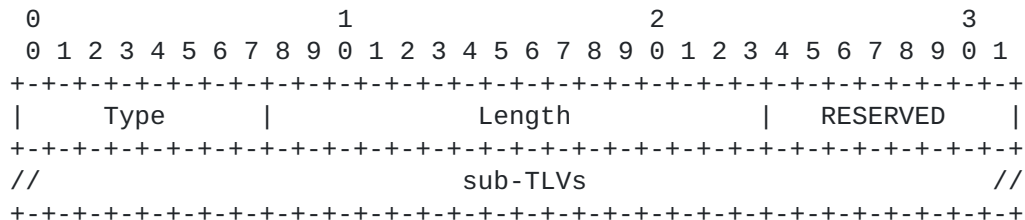
The Segment List sub-TLV encodes a single explicit path towards the endpoint as described in section 5.1 in [\[I-D.ietf-spring-segment-routing-policy\]](#). The Segment List sub-TLV includes the elements of the paths (i.e., segments) as well as an optional Weight sub-TLV.

The Segment List sub-TLV may exceed 255 bytes length due to large number of segments. Therefore a 2-octet length is required. According to [\[I-D.ietf-idr-tunnel-encaps\]](#), the first bit of the sub-TLV codepoint 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:



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.
- o RESERVED: 1 octet of reserved bits. SHOULD 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.

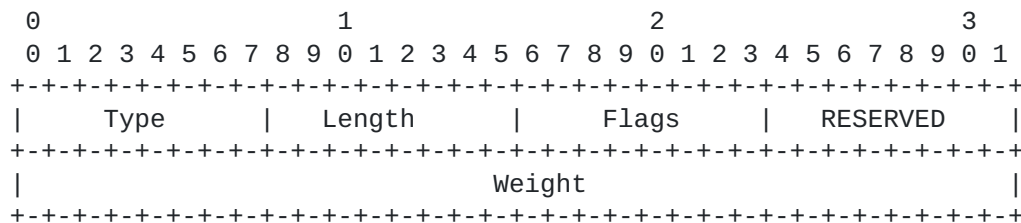
Validation of an explicit path encoded by the Segment List sub-TLV is beyond the scope of BGP and performed by the SRPM as described in section 5 in [[I-D.ietf-spring-segment-routing-policy](#)].

2.4.4.1. Weight Sub-TLV

The Weight sub-TLV specifies the weight associated to 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:



where:

- o Type: 9.
- o Length: 6
- o Flags: 1 octet of flags. None are defined at this stage. Flags SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o RESERVED: 1 octet of reserved bits. SHOULD be set to zero on transmission and MUST be ignored on receipt.

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.

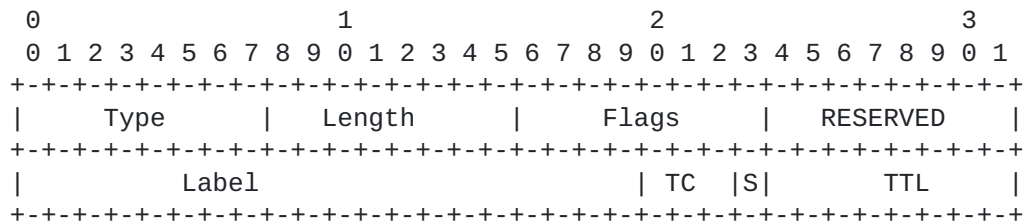
[I-D.ietf-spring-segment-routing-policy] defines several Segment Types:

- Type A: SID only, in the form of MPLS Label
- Type B: SID only, in the form of IPv6 address
- Type C: IPv4 Node Address with optional SID
- Type D: IPv6 Node Address with optional SID for SR MPLS
- Type E: IPv4 Address and index with optional SID
- Type F: IPv4 Local and Remote addresses with optional SID
- Type G: IPv6 Address and index for local and remote pair with optional SID for SR MPLS
- Type H: IPv6 Local and Remote addresses with optional SID for SR MPLS
- Type I: IPv6 Node Address with optional SID for SRv6
- Type J: IPv6 Address and index for local and remote pair with optional SID for SRv6
- Type K: IPv6 Local and Remote addresses for SRv6

The follow sub-sections specify the sub-TLV used for encoding each of these Segment Types.

2.4.4.2.1. Type A: SID only, in the form of MPLS Label

The Type A Segment Sub-TLV encodes a single SR-MPLS SID. The format is as follows:



where:

- o Type: 1.
- o Length is 6.
- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD be set to zero on transmission and MUST be ignored on receipt.

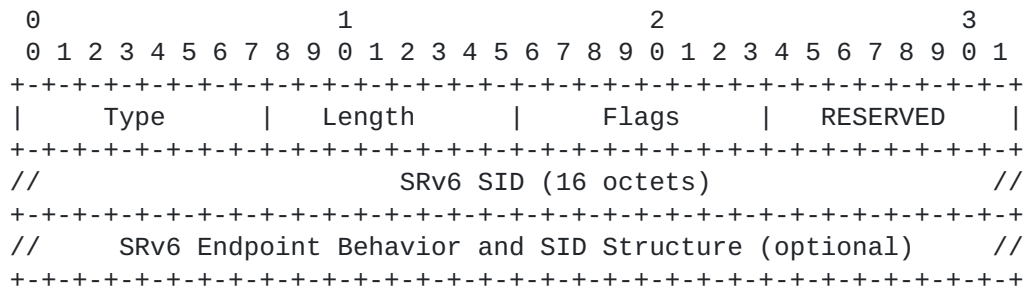
- o Label: 20 bits of label value.
- o TC: 3 bits of traffic class.
- o S: 1 bit of bottom-of-stack.
- o TTL: 1 octet of TTL.

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

- o The S bit SHOULD be zero upon transmission, and MUST be ignored upon reception.
- o If the originator wants the receiver to choose the TC value, it sets the TC field to zero.
- o If the originator wants the receiver to choose the TTL value, it sets the TTL field to 255.
- o If the originator wants to recommend a value for these fields, it puts those values in the TC and/or TTL fields.
- o 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. Type B: SID only, in the form of IPv6 address

The Type B Segment Sub-TLV encodes a single SRv6 SID. The format is as follows:



where:

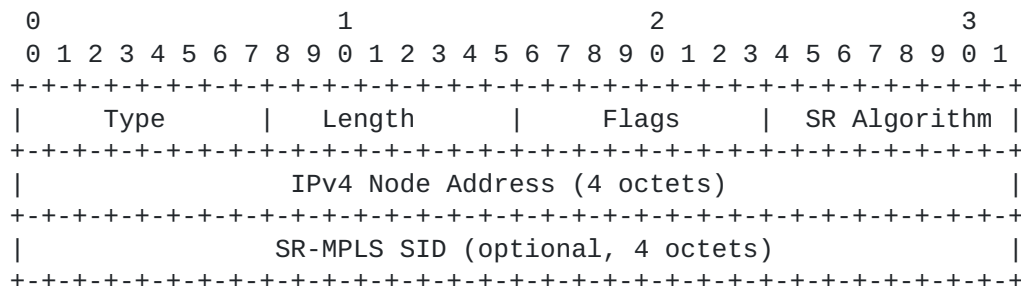
- o Type: 13.
- o Length is variable.

- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o SRV6 SID: 16 octets of IPv6 address.
- o SRV6 Endpoint Behavior and SID Structure: Optional, as defined in [Section 2.4.4.2.13](#).

The TLV 2 defined for advertisement of Segment Type B in the earlier versions of this document has been deprecated to avoid backward compatibility issues (refer [Appendix A](#) for details).

2.4.4.2.3. Type C: IPv4 Node Address with optional SID

The Type C Segment Sub-TLV encodes an IPv4 node address, SR Algorithm and an optional SR-MPLS SID. The format is as follows:



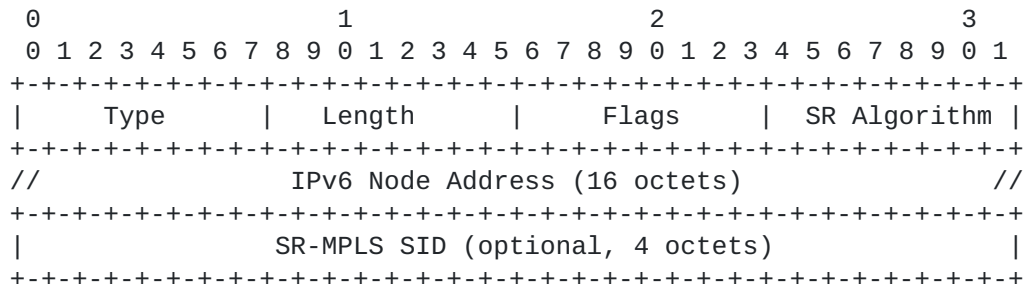
where:

- o Type: 3.
- o Length is 10 when the SR-MPLS SID is present else is 6.
- 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 SHOULD be set to zero on transmission and MUST be ignored on receipt.
- 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](#).

2.4.4.2.4. Type D: IPv6 Node Address with optional SID for SR MPLS

The Type D Segment Sub-TLV encodes an IPv6 node address, SR Algorithm and an optional SR-MPLS SID. The format is as follows:



where:

- o Type: 4
- o Length is 22 when the SR-MPLS SID is present else is 18.
- 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 SHOULD be set to zero on transmission and

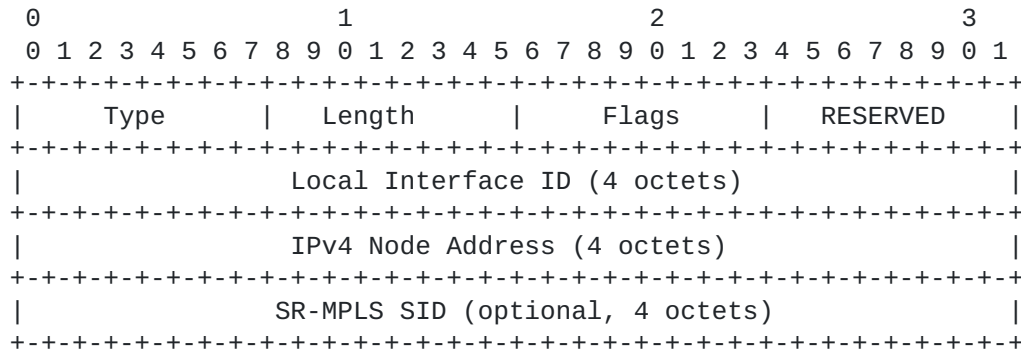
MUST

be ignored on receipt.

- o IPv6 Node Address: a 16 octet IPv6 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](#).

2.4.4.2.5. Type E: IPv4 Address + Local Interface ID with optional SID

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:

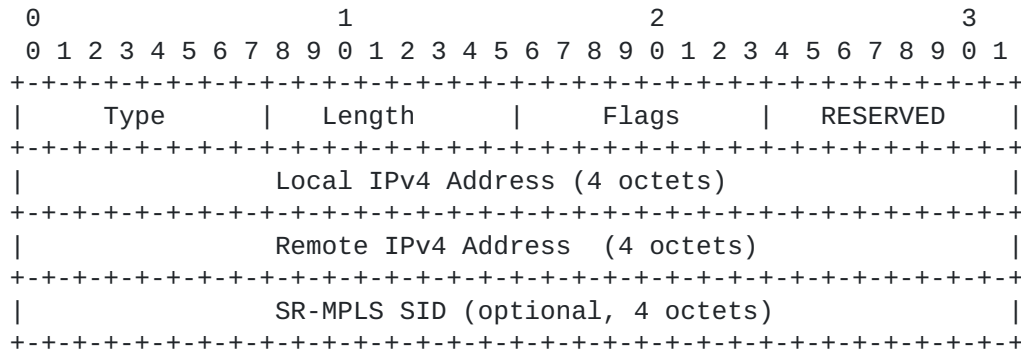


where:

- o Type: 5.
- o Length is 14 when the SR-MPLS SID is present else is 10.
- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD 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. Type F: IPv4 Local and Remote addresses with optional SID

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:

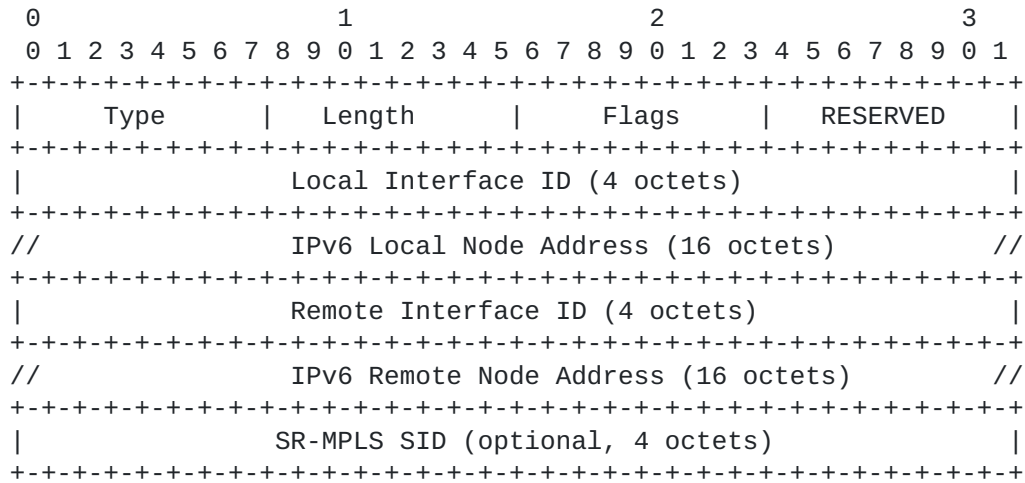


where:

- o Type: 6.
- o Length is 14 when the SR-MPLS SID is present else is 10.
- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o Local IPv4 Address: a 4 octet IPv4 address.
- o Remote IPv4 Address: a 4 octet IPv4 address.
- o SR-MPLS SID: optional, 4 octet field containing label, TC, S and TTL as defined in [Section 2.4.4.2.1](#).

2.4.4.2.7. Type G: IPv6 Address + Interface ID for local and remote pair with optional SID for SR MPLS

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:

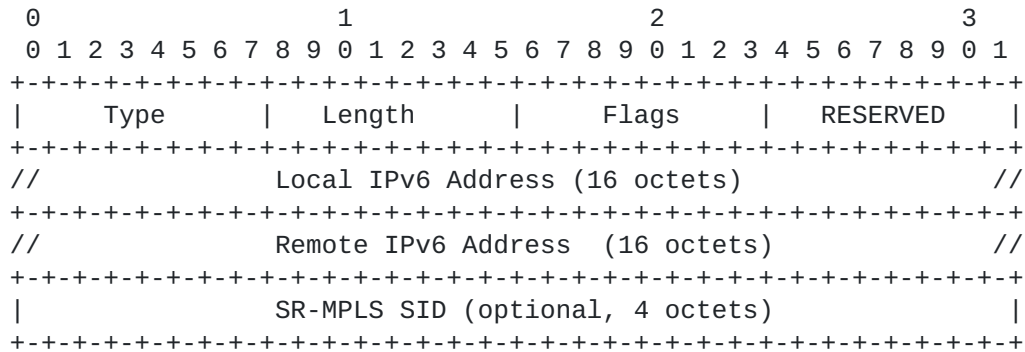


where:

- o Type: 7
- o Length is 46 when the SR-MPLS SID is present else is 42.
- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD 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. Type H: IPv6 Local and Remote addresses with optional SID for SR MPLS

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:



where:

- o Type: 8
- o Length is 38 when the SR-MPLS SID is present else is 34.
- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD 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. Type I: IPv6 Node Address with optional SRv6 SID

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) //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

- o Type: 14
- o Length is variable.
- 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 SHOULD be set to zero on transmission and

MUST

be ignored on receipt.

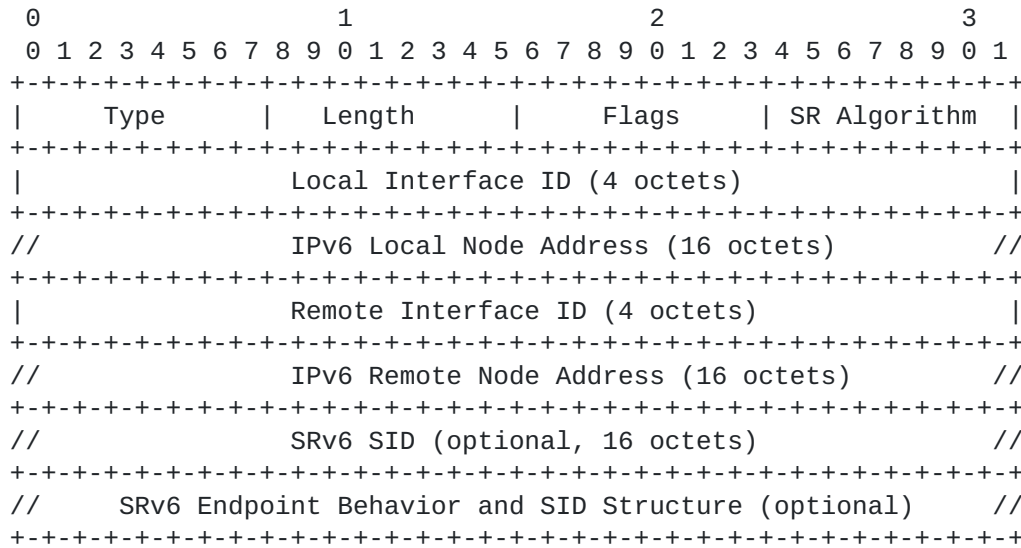
- o IPv6 Node Address: a 16 octet IPv6 address.
- o SRv6 SID: optional, 16 octet IPv6 address.
- o SRv6 Endpoint Behavior and SID Structure: Optional, as defined in [Section 2.4.4.2.13](#).

The TLV 10 defined for advertisement of Segment Type I in the earlier

versions of this document has been deprecated to avoid backward compatibility issues (refer [Appendix A](#) for details).

2.4.4.2.10. Type J: IPv6 Address + Interface ID for local and remote pair for SRv6 with optional SID

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:



where:

- o Type: 15
- o Length is variable.
- 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 SHOULD be set to zero on transmission and

MUST

be ignored on receipt.

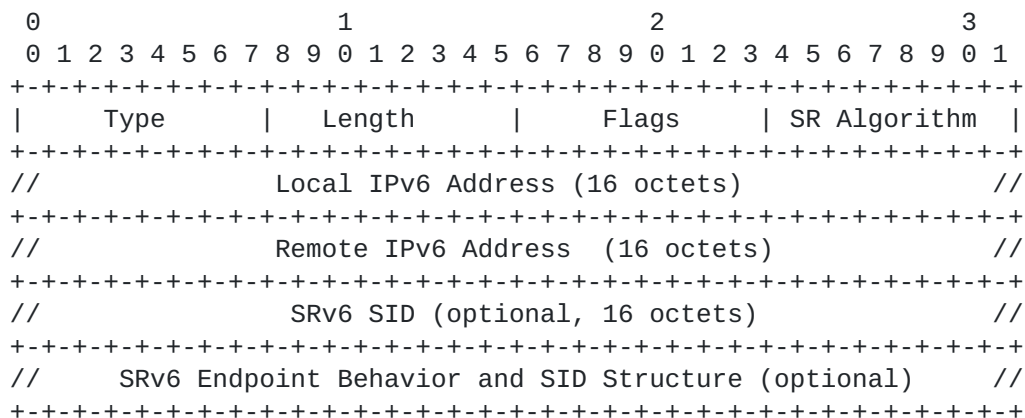
- o Local Interface ID: 4 octets of interface index as defined in [\[RFC8664\]](#).
- o IPv6 Local Node Address: a 16 octet IPv6 address.
- o Remote Interface ID: 4 octets of interface index as defined in [\[RFC8664\]](#). The value MAY be set to zero when the local node address and interface identifiers 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, 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 advertisement of Segment Type J in the earlier versions of this document has been deprecated to avoid backward compatibility issues (refer [Appendix A](#) for details).

2.4.4.2.11. Type K: IPv6 Local and Remote addresses for SRV6 with optional SID

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:



where:

- o Type: 16
- o Length is variable.
- 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 SHOULD 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 SRV6 SID: optional, 16 octet IPv6 address.
- o SRV6 Endpoint Behavior and SID Structure: Optional, as defined in [Section 2.4.4.2.13](#).

The TLV 12 defined for advertisement of Segment Type K in the earlier versions of this document has been deprecated to avoid backward compatibility issues (refer [Appendix A](#) for details).

2.4.4.2.12. Segment Flags

The Segment Types sub-TLVs described above MAY contain following flags in the "Flags" field defined in [Section 6.8](#):

```

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

```

where:

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

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 in [\[I-D.ietf-spring-segment-routing-policy\]](#).

MPLS S-Flag: This flag, when set, indicates the presence of the SR- 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](#).

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

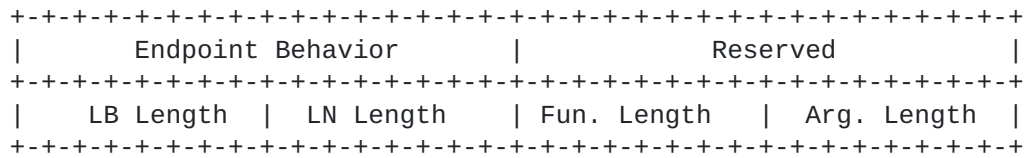
The following applies to the Segment Flags:

- o V-Flag is applicable to all Segment Types.

- o A-Flag is applicable 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.
- o S-Flag is applicable 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.
- o B-Flag is applicable 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 [[I-D.ietf-spring-srv6-network-programming](#)] encoding as described below:



where:

Endpoint Behavior: 2 octets. The SRv6 Endpoint Behavior code point for this SRv6 SID as defined in section 9.2 of [[I-D.ietf-spring-srv6-network-programming](#)]. When set with the value 0, the choice of SRv6 Endpoint Behavior is left to the headend.

Reserved: 2 octet of reserved bits. SHOULD 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

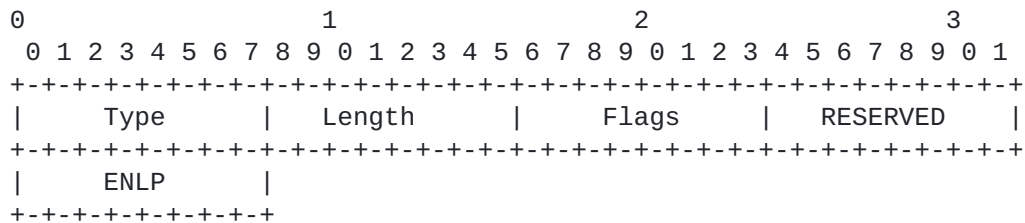
In order to steer an unlabeled IP packet into an SR policy, it is necessary to create a label stack for that packet, and to push one or more labels onto that stack.

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](#)].



Where:

Type: 14.

Length: 3.

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

RESERVED: 1 octet of reserved bits. SHOULD 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 [\[I-D.ietf-spring-segment-routing-policy\]](#) draft describes the behavior on the headend for handling of 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 2.11 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	2	3	4	5	6	7	8	9																				
Type															Length															Priority															RESERVED														

Where:

Type: 15

Length: 2.

Priority: a 1-octet value.

RESERVED: 1 octet of reserved bits. SHOULD 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 length due to long name. Therefore a 2-octet length is required.

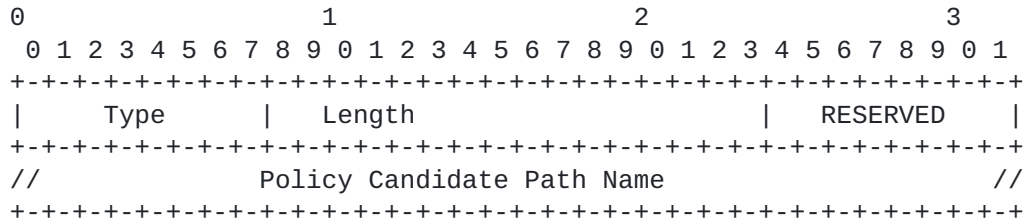
According

to [[I-D.ietf-idr-tunnel-encaps](#)], the first bit of the sub-TLV codepoint 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 be 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:



Where:

Type: 129.

Length: Variable.

RESERVED: 1 octet of reserved bits. SHOULD 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.

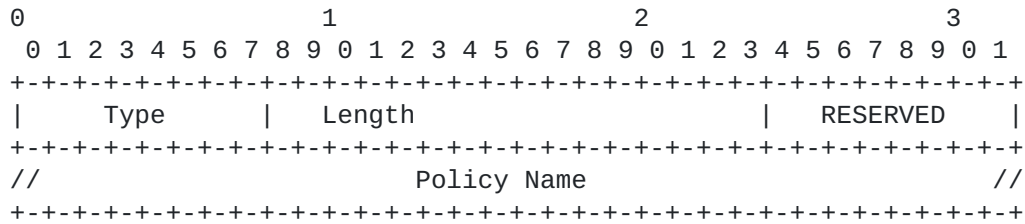
Usage of Policy Name sub-TLV is described in section 2.1 of [\[I-D.ietf-spring-segment-routing-policy\]](#).

The Policy Name sub-TLV may exceed 255 bytes length due to long policy name. Therefore a 2-octet length is required. According to [\[I-D.ietf-idr-tunnel-encaps\]](#), the first bit of the sub-TLV codepoint defines the size of the length field. Therefore, for the Policy Name

sub-TLV a code point of 128 or higher is used.

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:



Where:

Type: TBD

Length: Variable.

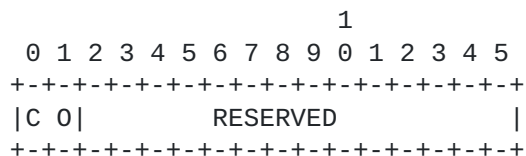
RESERVED: 1 octet of reserved bits. SHOULD be set to zero on transmission and MUST be ignored on receipt.

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 as defined in [\[I-D.ietf-idr-tunnel-encaps\]](#) is used to steer traffic into a policy.

When the Color Extended Community is used for the purpose of steering the traffic into an SR Policy, two bits from the Flags field (as defined in [\[I-D.ietf-idr-tunnel-encaps\]](#)) are used as follows:



where CO bits are defined as the "Color-Only" bits. [\[I-D.ietf-spring-segment-routing-policy\]](#) defines the influence of these bits on the automated steering of BGP Payload traffic onto SR Policies.

4. SR Policy Operations

As described in this document, 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 is 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 head-ends.

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-target SHOULD be attached to the advertisement, where each route-target identifies one or more intended head-ends 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 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's 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

A 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 it's becoming unusable, BGP MUST delete it's 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 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.

5. Error Handling

This section describes the error handling actions, as described in [\[RFC7606\]](#), that are to be performed for handling of 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

Previdi, et al.
34]

Expires May 17, 2021

[Page

validation of length of each NLRI and the total length of the MP_REACH_NLRI and MP_UNREACH_NLRI attributes.

When the error determined allows for the router to skip the malformed

NLRI(s) and continue 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 [[I-D.ietf-idr-tunnel-encaps](#)] 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 on the basis of such a validation.

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

6. IANA Considerations

This document requests codepoint allocations for new TLVs/sub-TLVs in following existing registries:

- o Subsequent Address Family Identifiers (SAFI) Parameters registry

o BGP Tunnel Encapsulation Attribute Tunnel Types registry under
the
BGP Parameters registry

Previdi, et al.
35]

Expires May 17, 2021

[Page

- o BGP Tunnel Encapsulation Attribute sub-TLVs registry under the BGP Parameters registry
- o Color Extended Community Flags registry under the BGP Extended Communities registry

This document also requests creation of the following new registries:

- o SR Policy Segment List Sub-TLVs under the BGP Parameters registry
- o SR Policy Binding SID Flags under the BGP Parameters registry
- o SR Policy Segment Flags under the BGP Parameters registry

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

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

Codepoint	Description	Reference
73	SR Policy SAFI	This document

6.2. Existing Registry: BGP Tunnel Encapsulation Attribute Tunnel Types

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

Codepoint	Description	Reference
15	SR Policy Type	This document

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

This document defines new sub-TLVs in the registry "BGP Tunnel Encapsulation Attribute sub-TLVs" that has been assigned codepoints by IANA as follows:

Codepoint	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
TBD	SRv6 Binding SID sub-TLV	This document
128	Segment List sub-TLV	This document
129	Policy CP Name sub-TLV	This document
TBD	Policy Name sub-TLV	This document

6.4. Existing Registry: Color Extended Community Flags

This document requests allocations in the registry called "Color Extended Community Flags" under the "BGP Tunnel Encapsulation Parameters" grouping.

Following bits are to be allocated:

Bit Position	Description	Reference
0-1	Color-only bits	This document

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

This document requests creation of a new registry called "SR Policy Segment List Sub-TLVs". The allocation policy of this registry is "Specification Required" according to [[RFC8126](#)].

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

Value	Description	Reference
1	Type A MPLS SID sub-TLV	This
2	Deprecated	This
3	Type C IPv4 Node and SID sub-TLV	This
4	Type D IPv6 Node and SID for SR-MPLS sub-TLV	This
5	Type E IPv4 Node, index and SID sub-TLV	This
6	Type F IPv4 Local/Remote addresses and SID sub-TLV	This
7	Type G IPv6 Node, index for remote and local pair and SID for SR-MPLS sub-TLV	This
8	Type H IPv6 Local/Remote addresses and SID sub-TLV	This
9	Weight sub-TLV	This
10	Deprecated	This
11	Deprecated	This
12	Deprecated	This
13	Type B SRv6 SID sub-TLV	This
14	Type I IPv6 Node and SID for SRv6 sub-TLV	This
15	Type J IPv6 Node, index for remote and local pair and SID for SRv6 sub-TLV	This
16	Type K IPv6 Local/Remote addresses and SID for SRv6 sub-TLV	This

6.6. New Registry: SR Policy Binding SID Flags

This document requests creation of a new registry called "SR Policy Binding SID Flags". The allocation policy of this registry is "Specification Required" according to [[RFC8126](#)].

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	

6.7. New Registry: SR Policy SRv6 Binding SID Flags

This document requests creation of a new registry called "SR Policy SRv6 Binding SID Flags". The allocation policy of this registry is "Specification Required" according to [[RFC8126](#)].

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	

6.8. New Registry: SR Policy Segment Flags

This document requests creation of a new registry called "SR Policy Segment Flags". The allocation policy of this registry is "Specification Required" according to [[RFC8126](#)].

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	

6.9. Guidance for Designated Experts

In all cases of review by the Designated Expert (DE) described here, the DE is expected to ascertain the existence of suitable documentation (a specification) as described in [[RFC8126](#)]. The DE is also expected to check the clarity of purpose and use of the requested code points. Additionally, the DE must verify that any request for one of these code points has been made available for review and comment within the IETF: the DE will post the request to the IDR Working Group mailing list (or a successor mailing list designated by the IESG). If the request comes from within the IETF, it should be documented in an Internet-Draft. Lastly, the DE must ensure that any other request for a code point does not conflict with work that is active or already published within the IETF.

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 the 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 AS/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 setup 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. 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 and Cheng Li for their comments and review of this document.

9. Contributors

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

10. References

10.1. Normative References

- [I-D.ietf-idr-tunnel-encaps]
Patel, K., Velde, G., Sangli, S., and J. Scudder, "The
BGP Tunnel Encapsulation Attribute", [draft-ietf-idr-tunnel-encaps-20](#) (work in progress), November 2020.
- [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-09](#) (work in progress), November 2020.
- [I-D.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., Voyer, D.,
Matsushima, S., and Z. Li, "SRv6 Network Programming",
[draft-ietf-spring-srv6-network-programming-24](#) (work in progress), October 2020.
- [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>>.
- [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>>.
- [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] Filss, 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>>.
- [RFC8664] Sivabalan, S., Filss, 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>>.

10.2. Informational References

- [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>>.

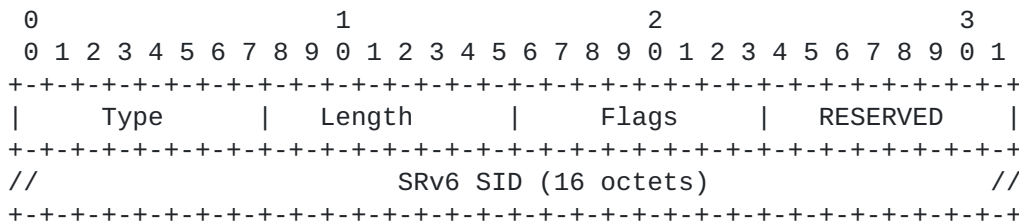
[RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", [RFC 6952](#), DOI 10.17487/RFC6952, May 2013, <<https://www.rfc-editor.org/info/rfc6952>>.

Appendix A. Deprecated Segment Sub-TLVs

This section describes the encoding of Segment Sub-TLVs that were specified in a previous version of this document and subsequently deprecated.

A.1. Type B-Deprecated: SID only, in the form of IPv6 address

The Type B-Deprecated Segment Sub-TLV encodes a single SRv6 SID. The format is as follows:

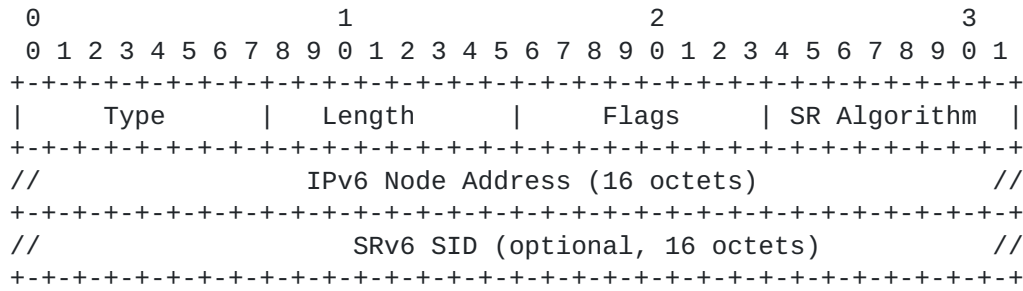


where:

- o Type: 2.
- o Length is 18.
- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o SRv6 SID: 16 octets of IPv6 address.

A.2. Type I-Deprecated: IPv6 Node Address with optional SRv6 SID

The Type I-Deprecated Segment Sub-TLV encodes an IPv6 node address, SR Algorithm and an optional SRv6 SID. The format is as follows:



where:

- o Type: 10.
- o Length is 34 when the SRv6 SID is present else is 18.
- 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 SHOULD be set to zero on transmission and

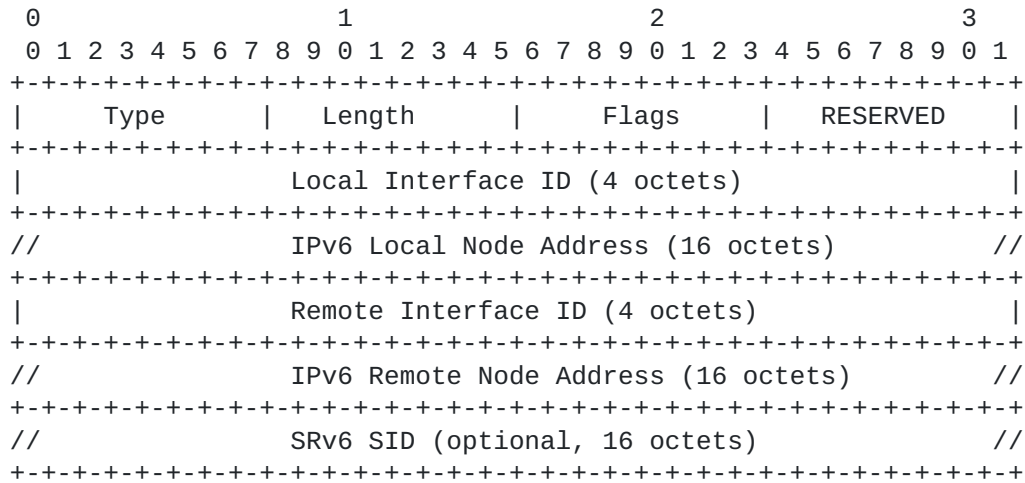
MUST

be ignored on receipt.

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

A.3. Type J-Deprecated: IPv6 Address + Interface ID for local and remote pair for SRv6 with optional SID

The Type J-Deprecated 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:

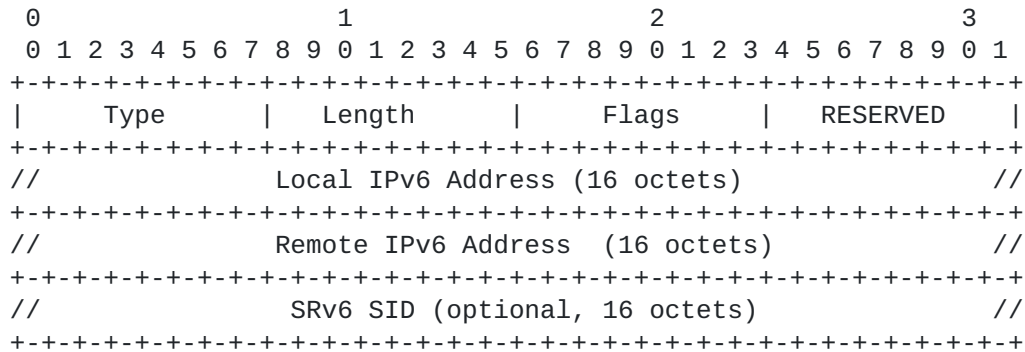


where:

- o Type: 11.
- o Length is 58 when the SRV6 SID is present else is 42.
- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD 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, 16 octet IPv6 address.

A.4. Type K-Deprecated: IPv6 Local and Remote addresses for SRv6 with optional SID

The Type K-Deprecated Segment Sub-TLV encodes an adjacency local address, an adjacency remote address and an optional SRv6 SID. The format is as follows:



where:

- o Type: 12.
- o Length is 50 when the SRv6 SID is present else is 34.
- o Flags: 1 octet of flags as defined in [Section 2.4.4.2.12](#).
- o RESERVED: 1 octet of reserved bits. SHOULD 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 SRv6 SID: optional, 16 octet IPv6 address.

Authors' Addresses

Stefano Previdi
Individual
IT

Email: stefano@previdi.net

Clarence Filsfils
Cisco Systems
Brussels
BE

Email: cfilsfil@cisco.com

Ketan Talaulikar (editor)
Cisco Systems
India

Email: ketant@cisco.com

Paul Mattes
Microsoft
One Microsoft Way
Redmond, WA 98052
USA

Email: pamattes@microsoft.com

Eric Rosen
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
US

Email: erosen@juniper.net

Dhanendra Jain
Google

Email: dhanendra.ietf@gmail.com

Steven Lin
Google

Email: stevenlin@google.com

