

Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: 30 November 2026

S. Ramasamy, Ed.
M. Sundaramoorthy
S. Sethuram
V. Kasiviswanathan
NextHop AI
29 May 2026

BGP Extensions for SRv6 Per-Flow Traffic Engineering: Composite
Candidate Path and Forwarding-Class Signaling
draft-ramasamy-idr-sr-policy-composite-perflow-00

Abstract

The Segment Routing Policy Architecture (RFC 9256) defines the concept of a Composite Candidate Path, which enables a parent SR Policy to recursively steer traffic into a set of constituent SR Policies. Section 8.6 of RFC 9256 further defines per-flow steering, in which packets are classified into a Forwarding Class value by a Quality-of-Service classifier at the headend, and that Forwarding Class value selects a specific constituent SR Policy, identified by its color, for forwarding.

RFC 9830 specifies how BGP distributes SR Policy Candidate Paths but explicitly excludes composite Candidate Paths from its scope. Consequently, no standard BGP mechanism currently exists to distribute a parent per-flow SR Policy -- including its Forwarding-Class-to-color mapping table -- to headend nodes. This document defines two new sub-TLVs for the BGP Tunnel Encapsulation Attribute (SR Policy type, Tunnel Type 15): the Constituent SR Policy sub-TLV and the nested Per-Flow Forwarding Class sub-TLV. Together, these extensions allow a controller to distribute a fully specified composite or per-flow Candidate Path to headend routers via BGP.

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 30 November 2026.

Copyright Notice

Copyright (c) 2026 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	3
1.1. Per-Flow Traffic Engineering	3
1.2. Scope	5
1.3. Requirements Language	5
2. Terminology	5
3. Problem Statement	7
3.1. Composite CPs Are Outside RFC 9830 Scope	7
3.2. Controller-to-Headend Signaling Gap	7
3.3. Interaction with the Color Extended Community	7
4. BGP SR Policy Encoding Extensions	8
4.1. Extended SR Policy Encoding Structure	8
4.2. Constituent SR Policy Sub-TLV	9
4.3. Per-Flow Forwarding Class (FC) Sub-TLV	11
5. Encoding Example	13
6. Procedures	14
6.1. Origination	14
6.2. Reception and SRPM Processing	15
6.3. Route Reflector Transparency	16
6.4. Graceful Restart	16
6.5. Capability Advertisement	16
6.6. Backward Compatibility	17
7. Error Handling	17
8. Relationship to Other Documents	19
8.1. RFC 9256 -- SR Policy Architecture	19
8.2. RFC 9830 -- Advertising SR Policies in BGP	20
8.3. BGP-LS Northbound Collection	20
8.4. PCEP Per-Flow Encoding	20
8.5. YANG Data Model	20
9. Operations	21
9.1. Distribution Ordering	21

9.2. Per-Peer State Limits	21
9.3. Convergence and Reconvergence	21
9.4. Observability	21
10. Security Considerations	22
11. IANA Considerations	23
11.1. BGP Tunnel Encapsulation Attribute Sub-TLVs Registry . .	23
11.2. Constituent SR Policy Sub-TLV Types Registry	24
12. Normative References	24
13. Informative References	25
Authors' Addresses	27

1. Introduction

Segment Routing (SR) [RFC8402] allows a headend node to steer packet flows along any explicitly specified path. The headend instantiates an SR Policy [RFC9256] identified by the tuple <Headend, Color, Endpoint>. A Candidate Path (CP) is the unit of signaling, and RFC 9830 specifies how BGP distributes these CPs to headends using a dedicated SR Policy SAFI (SAFI 73) with Tunnel Encapsulation Attribute sub-TLVs.

RFC 9256 Section 2.2 defines three types of Candidate Paths:

- * ***Explicit:** A specific set of Segment Lists.
- * ***Dynamic:** An optimization objective and constraints.
- * ***Composite:** A container grouping multiple constituent SR Policies for combined steering.

The Composite CP enables the construction of a parent SR Policy that distributes traffic across multiple constituent SR Policies, each identified by a distinct color and sharing the same headend and endpoint as the parent.

1.1. Per-Flow Traffic Engineering

Section 8.6 of RFC 9256 specifies per-flow steering. A headend maintains a forwarding array indexed by a Forwarding Class (FC) value. Each index maps to either:

- * The IGP (Interior Gateway Protocol) shortest path to the endpoint (the default), or
- * A constituent SR Policy identified by its color.

A Quality-of-Service (QoS) classifier on ingress assigns each packet an FC value derived from DSCP (Differentiated Services Code Point) markings, 802.1p bits, or other local policy. The headend then steers the packet into the SR Policy (or IGP path) corresponding to that FC.

The FC field width and value range are encoding choices of this document. This document encodes FC as a 3-bit value (range 0-7); see Section 4.3. Implementations that natively classify traffic into more than 8 forwarding classes are responsible for mapping their internal classes onto the 0-7 range before BGP signaling.

The configuration model below illustrates this construct (illustrative, implementation-agnostic syntax):

```
! Segment lists for constituent policies
segment-list SL_GOLD
  index 10 srv6-sid 2001:db8:100::1

segment-list SL_SILVER
  index 10 srv6-sid 2001:db8:200::1

! Child SR Policy: color 100, endpoint 2001:db8::1
policy color 100 endpoint 2001:db8::1
  candidate-path preference 100 name p100 explicit
  segment-list SL_GOLD

! Child SR Policy: color 200, same endpoint
policy color 200 endpoint 2001:db8::1
  candidate-path preference 100 name p200 explicit
  segment-list SL_SILVER

! Parent per-flow SR Policy: color 300, endpoint 2001:db8::1
!   FC 1 -> color 100 (high-priority path)
!   FC 2 -> color 200 (best-effort path)
!   FC 0,3-7 -> IGP (default)
policy color 300 endpoint 2001:db8::1
  candidate-path preference 70 name pf300 per-flow
  forwarding-class 1 color 100
  forwarding-class 2 color 200
  forwarding-class default action igp
```

In a controller-driven deployment, the controller must distribute this parent policy -- including the FC-to-color mapping -- to the headend via BGP. RFC 9830 provides no mechanism to do so. This document fills that gap.

1.2. Scope

This document:

- * Defines the *Constituent SR Policy sub-TLV* as a new optional sub-TLV within the BGP Tunnel Encapsulation Attribute SR Policy type (Tunnel Type 15).
- * Defines the *Per-Flow Forwarding Class (FC) sub-TLV* as a nested optional sub-TLV within the Constituent SR Policy sub-TLV.
- * Specifies procedures for origination, advertisement, reception, and SR Policy Module (SRPM) processing of composite and per-flow Candidate Paths distributed via BGP.
- * Specifies error handling for malformed sub-TLVs.
- * Requests IANA allocations for the new sub-TLV code points.

This document does not modify any sub-TLV defined in RFC 9830 or RFC 9012. The SR Policy BGP SAFI (Subsequent Address Family Identifier), NLRI format, RR propagation rules, and best-path selection procedures of RFC 9830 remain unchanged. Inter-AS distribution of Composite Candidate Paths is out of scope; the same operational caveats as RFC 9830 Section 8 apply. PCEP (Path Computation Element Communication Protocol)-based composite CP distribution is addressed by [I-D.ietf-pce-multipath] and is out of scope.

This document covers similar ground to other individual Internet-Drafts addressing BGP distribution of Composite Candidate Paths (e.g., [I-D.jiang-idr-sr-policy-composite-path]). The authors encourage coordination and possible consolidation with such efforts during IDR WG discussion.

1.3. 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. Terminology

This document uses terms defined in [RFC9256] and [RFC9830]. Additional terms used are:

SR Policy: Identified by <Headend, Color, Endpoint>; a source-

routing policy defined by an ordered list of segments [RFC9256].

Candidate Path (CP): The unit of signaling for an SR Policy. May be explicit, dynamic, or composite [RFC9256].

Composite CP: A CP that groups constituent SR Policies for combined or per-flow steering (RFC 9256, Section 2.2).

Constituent SR Policy: A child SR Policy forming part of a Composite CP; referenced solely by color (RFC 9256, Section 2.2).

FC (Forwarding Class): A value assigned to a packet by an ingress QoS classifier and used to select a per-flow forwarding entry, consistent with the Forwarding Class concept of RFC 9256, Section 8.6. This document encodes FC as a 3-bit value (range 0-7); see Section 4.3.

Per-Flow Steering: A mode where each FC maps to a constituent SR Policy or the IGP path, consistent with RFC 9256, Section 8.6.

Parent SR Policy: The SR Policy whose active CP is a Composite CP. Has its own distinct color.

SRPM: SR Policy Module -- the headend block that processes SR Policy information and installs it in the forwarding plane [RFC9830].

BFD: Bidirectional Forwarding Detection [RFC5880].

DSCP: Differentiated Services Code Point [RFC2474].

ENLP: Explicit NULL Label Policy sub-TLV [RFC9830].

EVPN: Ethernet Virtual Private Network [RFC7432].

IGP: Interior Gateway Protocol.

PCE: Path Computation Element.

PCEP: Path Computation Element Communication Protocol.

QoS: Quality of Service.

ROV: Route Origin Validation [RFC6811].

SAFI: Subsequent Address Family Identifier.

VPN: Virtual Private Network.

3. Problem Statement

3.1. Composite CPs Are Outside RFC 9830 Scope

RFC 9830 Section 1 explicitly states: `"The signaling of Dynamic and Composite CPs (Sections 5.2 and 5.3, respectively, of [RFC9256]) is outside the scope of this document."` As a consequence, RFC 9830 defines no sub-TLV encoding for: the list of constituent SR Policies forming a Composite CP; the FC-to-color mapping enabling per-flow steering; or the default action for unmatched FC values.

3.2. Controller-to-Headend Signaling Gap

In a controller-driven SRv6 TE deployment, an SR Policy Controller distributes SR Policies to headends via BGP SAFI 73. For per-flow TE:

1. The controller computes constituent SR Policies (e.g., color 100 for high-priority traffic, color 200 for best-effort).
2. The controller computes the parent per-flow SR Policy (e.g., color 300) with the FC mapping.
3. The controller distributes constituent policies (colors 100 and 200) via RFC 9830. This works today.
4. The controller distributes the parent per-flow policy (color 300) via BGP. This fails -- no composite/FC encoding exists in RFC 9830.

Step 4 forces operators to configure the parent policy via CLI or NETCONF on every headend, defeating controller-driven automation. This document addresses that gap.

3.3. Interaction with the Color Extended Community

The Color Extended Community [RFC9012] enables BGP service routes (unicast, VPN, EVPN) to be steered into SR Policies at the headend by matching the community color to the SR Policy color. For per-flow TE:

- * Service routes carry the color of the `*parent*` SR Policy (e.g., color 300).
- * The headend steers matching traffic into the parent policy.
- * The parent policy's per-flow Composite CP steers each packet to a constituent SR Policy based on its FC value.

The Color Extended Community mechanism requires no change. The extensions defined in this document operate entirely at the SR Policy SAFI layer and do not affect service route processing.

4. BGP SR Policy Encoding Extensions

4.1. Extended SR Policy Encoding Structure

RFC 9830 defines the SR Policy encoding using the BGP Tunnel Encapsulation Attribute (Attribute Type 23, Tunnel Type 15). This document extends that structure with the Constituent SR Policy sub-TLV as a new optional sub-TLV within Tunnel Type 15.

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

Attributes:

Tunnel Encapsulation Attribute (Type 23)

Tunnel Type: SR Policy (15)

Binding SID sub-TLV	[optional, RFC 9830]
SRv6 Binding SID sub-TLV	[optional, RFC 9830]
Preference sub-TLV	[optional, RFC 9830]
Priority sub-TLV	[optional, RFC 9830]
SR Policy Name sub-TLV	[optional, RFC 9830]
SR Policy CP Name sub-TLV	[optional, RFC 9830]
ENLP sub-TLV	[optional, RFC 9830]
Constituent SR Policy sub-TLV	[optional, THIS DOCUMENT]
RESERVED (2 octets)	
Color (4 octets)	
Weight (4 octets)	
Per-Flow FC sub-TLV	[optional, THIS DOCUMENT]
Constituent SR Policy sub-TLV	[optional, MAY repeat]
...	

***Mutual Exclusion:** A Constituent SR Policy sub-TLV and a Segment List sub-TLV (Type 128, RFC 9830) MUST NOT appear in the same SR Policy Tunnel Type encoding. A Candidate Path is either composite (uses Constituent SR Policy sub-TLVs) or explicit/dynamic (uses Segment List sub-TLVs). Presence of both is a malformed condition (Section 7, Condition 1).

***Candidate-path selection:** When both Composite and explicit/dynamic Candidate Paths are advertised for the same <Headend, Color, Endpoint>, Candidate Path selection follows Section 2.9 of [RFC9256]. The composite encoding does not alter the preference-based ordering.

4.2. Constituent SR Policy Sub-TLV

The Constituent SR Policy sub-TLV is a new optional sub-TLV of the BGP Tunnel Encapsulation Attribute SR Policy Tunnel Type. It encodes a single constituent SR Policy within a Composite CP. Multiple occurrences MAY appear in the same SR Policy encoding, each representing one constituent SR Policy. The ordering of occurrences is not significant.

Wire 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										RESERVED																			
										Color																													
										Weight																													
										sub-TLVs (variable length, optional)																													

Fields:

Type: To be assigned by IANA from the "BGP Tunnel Encapsulation Attribute Sub-TLVs" registry under the Standards Action policy.

Length: 1-octet unsigned integer. Total length in octets of all fields following the Length field (RESERVED + Color + Weight + sub-TLVs). The minimum value is *10* (2 octets RESERVED + 4 octets Color + 4 octets Weight). A Length value less than 10 is a malformed condition (Section 7, Condition 8).

RESERVED: 2 octets. MUST be set to zero on transmission and MUST be ignored on receipt. The 2-octet RESERVED field preserves 32-bit word alignment for the Color and Weight fields.

The 1-octet Length field limits the total nested sub-TLV content per Constituent SR Policy sub-TLV to 245 octets (255 minus the 10 bytes of fixed fields). Future nested sub-TLVs requiring more than 245 octets of combined content will require a separate 2-octet-Length Constituent sub-TLV code point to be defined at that time.

Color: 4-octet unsigned non-zero integer. Identifies the

constituent SR Policy. Constituent SR Policies share the headend and endpoint of the parent SR Policy; they are uniquely identified by this Color within the context of the parent. The Color MUST NOT equal zero (Section 7, Condition 9) and MUST NOT equal the Color of the parent SR Policy in the NLRI (Section 7, Condition 7). All Constituent SR Policy sub-TLVs within a single SR Policy encoding MUST carry distinct Color values (Section 7, Condition 10).

Weight: 4-octet unsigned integer. Specifies the weight of this constituent SR Policy for weighted load-balancing, consistent with Section 2.11 of [RFC9256]. The Weight field is always present on the wire as a fixed 4-octet field; there is no implicit/absent default on receipt. On origination, controllers SHOULD encode Weight=1 when no explicit weight policy is configured.

In **weighted load-balancing mode** (no Per-Flow FC sub-TLV present in any Constituent SR Policy sub-TLV in the encoding), a Weight value of zero marks that constituent as receiving no traffic; this is a constituent-level validity outcome, not a treat-as-withdraw condition. If the Weight values of *_all_* constituents in the encoding are zero, the Composite CP carries no traffic and is treated as malformed (Section 7, Condition 3).

In **per-flow steering mode** (every Constituent SR Policy sub-TLV carries a Per-Flow FC sub-TLV), the Weight field has no load-balancing effect between FC-mapped constituents (per-flow steering directs each classified packet to a single constituent) and MUST be ignored on receipt. Originators MAY encode any value; the value is not interpreted. Weighted load balancing within the active CP of the constituent SR Policy itself remains governed by Section 2.11 of [RFC9256].

sub-TLVs: Variable-length field containing zero or one Per-Flow FC sub-TLV (Section 4.3) and optionally other sub-TLVs defined in the future.

Semantics -- Weighted Load Balancing Mode: When no Per-Flow FC sub-TLV is present in a Constituent SR Policy sub-TLV, the constituent participates in weighted load balancing over all traffic steered into the parent policy. The traffic fraction steered to this constituent is $w_i / \sum_j(w_j)$, where w_i is this Weight and the sum is over all valid constituents in the Composite CP.

Semantics -- Per-Flow Steering Mode: When a Per-Flow FC sub-TLV is present, the constituent receives only traffic classified to the specified FC value.

Mixed encoding is not permitted: all Constituent SR Policy sub-TLVs in a single encoding MUST either all include a Per-Flow FC sub-TLV or all omit it. The rationale is that mixing FC-classified and weighted-LB traffic in one Composite CP would require defining a residual-traffic policy for unclassified FC values in the presence of weighted constituents, which is not specified and creates ambiguous behavior. A mixed encoding is a malformed condition (Section 7, Condition 6).

4.3. Per-Flow Forwarding Class (FC) Sub-TLV

The Per-Flow FC sub-TLV is a nested optional sub-TLV within the Constituent SR Policy sub-TLV. It assigns a 3-bit FC value to the containing constituent SR Policy, enabling QoS-based per-flow steering.

Wire Format:

```

      0                               1
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+---+---+---+---+---+---+---+---+---+
|      Type      | Length(=2) |
+---+---+---+---+---+---+---+---+
|D|      RESERVED      | FC |
+---+---+---+---+---+---+---+---+

```

The sub-TLV is 4 octets total: 1-octet Type + 1-octet Length + 2 octets of value content.

The 2-octet value field is laid out as follows, where bit 0 is the most significant bit of the first value octet:

```

Bit 0      : D (Default-Drop flag)
Bits 1-12   : RESERVED (12 bits)
Bits 13-15  : FC (3 bits; LSBs of the second value octet)

```

All 16 bits of the value field are accounted for: 1 (D) + 12 (RESERVED) + 3 (FC) = 16 bits.

Fields:

Type: To be assigned by IANA from the new "Constituent SR Policy Sub-TLV Types" registry (Section 11.2).

Length: 1 octet. MUST be 2. A value other than 2 is a malformed condition (Section 7, Condition 5).

D (Default-Drop flag, bit 0): When set to 1, if the constituent SR

Policy referenced by this sub-TLV becomes invalid, traffic for this FC entry MUST be dropped rather than forwarded via the IGP path. When set to 0 (default), traffic falls back to the IGP shortest path to the endpoint. This flag provides an extension point for the default-action choice described in RFC 9256, Section 8.6; absent this flag the per-flow default action is resolved via local policy per that section.

RESERVED (bits 1-12, 12 bits): MUST be set to zero on transmission and MUST be ignored on receipt.

FC (bits 13-15, 3 bits): The Forwarding Class value for this constituent SR Policy. Valid range: 0-7. Different Constituent SR Policy sub-TLVs within the same SR Policy encoding MUST carry distinct FC values (Section 7, Condition 2).

Nested-registry Length convention. All sub-TLVs registered in the "Constituent SR Policy Sub-TLV Types" registry (Section 11.2) -- across both the Standards Action range (1-127) and the Expert Review range (128-254) -- use a *1-octet Length field*, regardless of Type value. This nested registry does not follow the RFC 9012, Section 2 high-bit Type / Length-width rule used for the parent BGP Tunnel Encapsulation Attribute Sub-TLVs registry. The 1-octet Length field bounds any single nested sub-TLV's value content to 255 octets, which is sufficient for the per-Constituent extension points anticipated by this document. Definitions of future nested sub-TLVs in this registry MUST conform to this convention.

Semantics: Upon receiving a Composite CP with Per-Flow FC sub-TLVs, the SRPM builds an FC-indexed steering array as follows:

- * All eight FC entries (0-7) are initialized to the IGP shortest path to the endpoint (the default action, consistent with RFC 9256, Section 8.6).
- * For each Constituent SR Policy sub-TLV carrying a Per-Flow FC sub-TLV, the array entry at index FC is set to the constituent SR Policy identified by Color.
- * If the constituent SR Policy for a given FC entry becomes invalid, the SRPM reverts that entry to the IGP shortest path (D=0) or drops traffic (D=1), consistent with RFC 9256, Section 8.6.
- * FC values not mapped by any Constituent SR Policy sub-TLV retain the default IGP action throughout the lifetime of the Composite CP.

5. Encoding Example

The following example encodes the parent per-flow SR Policy from Section 1.1 (color 300, endpoint 2001:db8::1, preference 70, CP name "pf300"). FC 1 maps to color 100 and FC 2 maps to color 200. All other FCs default to IGP (D=0 for both).

The Weight value of 1 shown in each Constituent SR Policy sub-TLV below is illustrative; because every constituent carries a Per-Flow FC sub-TLV, this is per-flow steering mode and the Weight field is ignored on receipt (Section 4.2).

SR Policy SAFI NLRI:

Distinguisher : 1
Color : 300
Endpoint : 2001:db8::1 (AFI=2, SAFI=73)

BGP Tunnel Encapsulation Attribute (Type 23, Tunnel Type 15):

Preference sub-TLV (Type 12):
Preference: 70

SR Policy CP Name sub-TLV (Type 129):
Name: "pf300"

Constituent SR Policy sub-TLV (Type TBA1):
Length: 14 (2 RESERVED + 4 Color + 4 Weight + 4 FC sub-TLV)
RESERVED: 0x0000
Color: 100
Weight: 1
Per-Flow FC sub-TLV (Type TBA2):
Length: 2
Value octet 0: 0x00 (D=0, RESERVED bits 1-7 = 0)
Value octet 1: 0x01 (bits 8-12 = 0; FC bits 13-15 = 001 = 1)
Decoded: D=0, RESERVED=0, FC=1 (FC 1 -> color 100)

Constituent SR Policy sub-TLV (Type TBA1):
Length: 14
RESERVED: 0x0000
Color: 200
Weight: 1
Per-Flow FC sub-TLV (Type TBA2):
Length: 2
Value octet 0: 0x00 (D=0, RESERVED bits 1-7 = 0)
Value octet 1: 0x02 (bits 8-12 = 0; FC bits 13-15 = 010 = 2)
Decoded: D=0, RESERVED=0, FC=2 (FC 2 -> color 200)

The SRPM on the receiving headend installs the following FC steering array for SR Policy (color 300, endpoint 2001:db8::1):

FC	Forwarding Action	Fallback on Invalid
0	IGP to 2001:db8::1 (default)	--
1	SR Policy color 100 -- SID 2001:db8:100::1	IGP (D=0)
2	SR Policy color 200 -- SID 2001:db8:200::1	IGP (D=0)
3	IGP to 2001:db8::1 (default)	--
4	IGP to 2001:db8::1 (default)	--
5	IGP to 2001:db8::1 (default)	--
6	IGP to 2001:db8::1 (default)	--
7	IGP to 2001:db8::1 (default)	--

Table 1

6. Procedures

6.1. Origination

An SR Policy Controller or BGP egress router originating a Composite or per-flow SR Policy CP MUST:

1. Set the SR Policy SAFI NLRI to represent the *parent* SR Policy (e.g., color 300, endpoint 2001:db8::1).
2. Include one Constituent SR Policy sub-TLV per constituent SR Policy. Each MUST carry a non-zero Color field that does not equal the parent policy's color or the color of any other Constituent SR Policy sub-TLV in the same encoding.
3. For per-flow steering, include a Per-Flow FC sub-TLV within each applicable Constituent SR Policy sub-TLV. All constituents MUST include the FC sub-TLV (no mixed encoding).
4. MUST NOT include a Segment List sub-TLV (Type 128) in the same encoding as any Constituent SR Policy sub-TLV.

5. Originate separate BGP SR Policy SAFI updates for each constituent (child) SR Policy using standard RFC 9830 Segment List sub-TLVs, as these are distinct NLRIs.
6. Distribute constituent (child) SR Policies **before** the parent Composite CP. While the SRPM tolerates transient inversion by treating unknown constituents as invalid (Section 6.2), controllers **SHOULD** sequence updates in child-then-parent order to minimize disruption during initial convergence.

6.2. Reception and SRPM Processing

Upon receiving a BGP UPDATE carrying an SR Policy SAFI NLRI with one or more Constituent SR Policy sub-TLVs, the BGP speaker applies standard BGP best-path selection per RFC 9830 to choose the best-path NLRI for the <Distinguisher, Color, Endpoint> tuple. The resulting Candidate Path is then offered to the SRPM along with locally configured and PCEP-signaled Candidate Paths for the same <Headend, Color, Endpoint> SR Policy. The SRPM performs its own Candidate Path selection across all sources per Section 2.9 of [RFC9256]; the BGP best-path output is one input to that selection, not the final outcome.

The SRPM **SHALL**:

1. Recognize the CP as a **Composite CP** type per RFC 9256, Section 2.2.
2. For each Constituent SR Policy sub-TLV, identify the corresponding child SR Policy at the local headend by matching <Headend, Color=sub-TLV Color, Endpoint>.
3. If a referenced constituent SR Policy is not locally known or has no valid CP, treat that constituent as **invalid**. The Composite CP remains valid as long as at least one constituent is valid, consistent with RFC 9256, Section 5.3. If all constituents are invalid, the Composite CP is invalid. If no other CP exists for the parent SR Policy, the parent SR Policy is invalid, and service routes steered by the Color Extended Community fall back to their unsteered BGP next hop.
4. Install the FC-indexed steering array per Section 4.3 semantics.
5. The SRPM **SHOULD** install the FC array **atomically** with respect to packet classification to avoid transient mis-steering during Composite CP changes.

6. Monitor liveness of constituent SR Policies. If a constituent transitions from valid to invalid, the SRPM MUST update the corresponding FC entry to the appropriate fallback (IGP path or drop per D flag) without waiting for a BGP update.

**Convergence behavior:* If the parent Composite CP arrives before one or more constituent SR Policies have been installed by BGP, the SRPM MUST treat those missing constituents as invalid (applying IGP or drop fallback for their FC entries) and MUST update the FC array when the constituent later becomes valid. Liveness monitoring (step 6) handles the transition once convergence completes.

6.3. Route Reflector Transparency

Route Reflectors (RRs) [RFC4456] that do not implement the sub-TLVs defined in this document MUST propagate the Constituent SR Policy sub-TLV and its nested Per-Flow FC sub-TLV unchanged, consistent with the handling of unknown sub-TLVs specified in Section 2 of [RFC9012], as applied by [RFC9830].

6.4. Graceful Restart

During BGP Graceful Restart [RFC4724] and Long-Lived Graceful Restart [RFC9494], stale Composite CP routes remain installed at the SRPM for the duration of the stale-routes timer. Constituent SR Policy liveness is handled independently per RFC 4724 / RFC 9494 for those NLRIs. The Composite CP need not be re-validated against constituent liveness during the stale interval, to avoid blackholing traffic during BGP session restart.

6.5. Capability Advertisement

A BGP speaker indicates support for Composite CP processing implicitly by advertising SR Policy SAFI (SAFI 73) capability in the BGP OPEN message. A headend that supports SAFI 73 but does not implement this document's extensions will ignore the Constituent SR Policy sub-TLV as an unknown sub-TLV per RFC 9830, rendering the parent CP invalid (Section 6.6). This is a safe failure mode. No separate capability code is defined.

Operators SHOULD verify that headend software supports Composite CP processing before distributing parent per-flow policy advertisements. Where available, the headend's YANG operational state model (see Section 8.5 for companion YANG status) or BGP-LS state collection (Section 8.3) may be used to confirm support; absent such telemetry, operators may rely on vendor release notes or controlled-rollout testing.

6.6. Backward Compatibility

BGP speakers that do not implement this document will encounter an unknown sub-TLV type in the SR Policy Tunnel Encapsulation. Per RFC 9830, unknown sub-TLVs are ignored. The SRPM will see no Segment List sub-TLVs and no Constituent SR Policy sub-TLVs, treating the Candidate Path as having no valid forwarding information and marking it invalid. No incorrect forwarding state is installed. This is a safe failure mode.

7. Error Handling

Error handling for BGP UPDATE messages follows the "treat-as-withdraw" strategy of [RFC7606] and RFC 9830, Section 5. Each of the following conditions constitutes a malformed SR Policy NLRI; the treat-as-withdraw procedure MUST be applied:

#	Malformed Condition
1	A Segment List sub-TLV (Type 128) and one or more Constituent SR Policy sub-TLVs co-exist in the same SR Policy encoding.
2	Two or more Constituent SR Policy sub-TLVs carry Per-Flow FC sub-TLVs with duplicate FC values.
3	In weighted load-balancing mode, the Weight values of ALL Constituent SR Policy sub-TLVs in the encoding are zero (so no traffic can be steered into any constituent). (Per-constituent Weight=0 is a constituent-level validity outcome -- see Section 4.2 -- and is not by itself a malformed NLRI. In per-flow steering mode, Weight is ignored entirely.)
4	A Per-Flow FC sub-TLV appears more than once within a single Constituent SR Policy sub-TLV.
5	A Per-Flow FC sub-TLV has a Length value other than 2.
6	Mixed encoding: some Constituent SR Policy sub-TLVs carry a Per-Flow FC sub-TLV and others do not.
7	A Constituent SR Policy sub-TLV carries a Color value equal to the parent SR Policy's Color in the NLRI.
8	A Constituent SR Policy sub-TLV carries a Length value less than 10.
9	A Constituent SR Policy sub-TLV carries a Color value of zero.
10	Two or more Constituent SR Policy sub-TLVs within the same SR Policy encoding carry identical Color values.
11	An SR Policy Tunnel Type encoding contains neither a Segment List sub-TLV nor any Constituent SR Policy sub-TLV (i.e., an empty Candidate Path with no forwarding contents).

Table 2

Cycle detection. Constituent SR Policies SHOULD reference non-composite SR Policies (i.e., the active CP of each constituent SHOULD itself be an explicit or dynamic CP, not a Composite CP). If transitive composition is permitted by local policy, the SRPM MUST bound recursion to a configurable maximum depth (default 2) and treat references that exceed that depth, or that form a transitive reference cycle, as ***invalid at the constituent level***: the affected constituent entry is marked invalid, the Composite CP remains valid if at least one other constituent is valid, and standard treat-as-withdraw does not apply. A directly self-referential Color (Color equals the parent Color in the NLRI) is already a malformed NLRI per Condition 7.

A Per-Flow FC sub-TLV is defined only inside a Constituent SR Policy sub-TLV (Section 4.3). The nested type space (Section 11.2) is independent of the top-level BGP Tunnel Encapsulation Attribute Sub-TLVs registry; a top-level Type value is always interpreted per the top-level registry, so an "FC sub-TLV at the top level" is not a detectable error condition and is therefore not enumerated here.

8. Relationship to Other Documents

8.1. RFC 9256 -- SR Policy Architecture

RFC 9256 is the architectural authority. Sections 2.2 and 8.6 define Composite CPs and per-flow steering respectively. This document implements the BGP signaling plane for those constructs. One gap remains in RFC 9256: the default action drop behavior is mentioned in some implementations but not normatively defined in RFC 9256, Section 8.6. The D flag in the Per-Flow FC sub-TLV (Section 4.3) provides a forward-compatible extension point for this choice.

SPRING WG coordination required. The normative semantics of the per-flow default-action -- and any future widening of the Forwarding Class field beyond 3 bits -- are architectural decisions that belong to the SPRING WG (responsible for RFC 9256). Adoption and progression of this document by the IDR WG SHOULD include explicit confirmation from the SPRING WG that (a) the 3-bit FC encoding adopted here is compatible with the RFC 9256 architecture, and (b) the D-flag semantics described in Section 4.3 are consistent with any future SPRING work on the default-action choice.

8.2. RFC 9830 -- Advertising SR Policies in BGP

RFC 9830 is the parent document extended here. All SAFI 73 operational procedures remain unchanged. This document does not carry an Updates: RFC 9830 header because it adds new optional sub-TLVs within the extensibility model that RFC 9830 already provides and does not modify any existing RFC 9830 normative behavior.

Note for Last Call review. RFC 9830, Section 1 explicitly scopes Composite Candidate Path signaling out of that document. A reviewer may argue that bringing Composite CPs into scope warrants an Updates: clause. The authors take the position that scope is set per-document and does not constrain extensibility, and that adding new sub-TLVs through RFC 9830's pre-existing extensibility model does not require an Updates: header. The shepherd writeup should record this position so the IESG/IDR review can resolve it explicitly rather than re-litigating it on the list.

8.3. BGP-LS Northbound Collection

A companion BGP-LS mechanism is expected to extend RFC 9857 with TLVs corresponding to the Constituent SR Policy and Per-Flow FC sub-TLVs defined here, enabling northbound collection of Composite CP state from headends to a controller. Implementations SHOULD support both southbound signaling (this document) and northbound state collection for full controller visibility.

8.4. PCEP Per-Flow Encoding

PCEP-based distribution of SR Policy Candidate Paths is defined in [RFC9862]. Composite and per-flow extensions to PCEP are being defined separately in the PCE Working Group ([I-D.ietf-pce-multipath]). The exact TLV names and object placement in that draft should be cross-checked against its latest revision before publication; this document does not normatively depend on those names. The BGP encoding in this document and the PCEP encoding are complementary; SRPM behavior on receipt is identical regardless of the signaling protocol used.

8.5. YANG Data Model

The SPRING WG YANG data model for SR Policy ([I-D.ietf-spring-sr-policy-yang]) currently excludes Composite CP modeling. A companion contribution to that draft is expected to add a composite CP type container, a constituent-sr-policy list (color, weight, fc-value, default-drop), and a per-flow-forwarding-class leaf-list. For current observability, BGP-LS [RFC9857] collection of Composite CP state (Section 8.3) is recommended.

9. Operations

9.1. Distribution Ordering

BGP provides no atomicity guarantee across separate NLRI updates. Controllers SHOULD distribute constituent (child) SR Policy CPs before the parent Composite CP, as specified in Section 6.1. Headends MUST tolerate transient inversions gracefully per Section 6.2, applying IGP or drop fallback for unresolved constituents until convergence completes.

9.2. Per-Peer State Limits

A controller advertising large numbers of Composite CPs referencing nonexistent or invalid constituents will cause the SRPM to maintain per-constituent invalid state and IGP-fallback entries. Implementations SHOULD enforce configurable per-peer limits on the number of Composite CP NLRIs and the number of Constituent SR Policy sub-TLVs per CP. Note that per-flow steering naturally caps the per-CP constituent count at 8 (one per FC value), but weighted load-balancing mode does not impose a comparable cap and is therefore the primary state-amplification surface. Exceeding configured limits SHOULD be logged and MAY result in treat-as-withdraw of excess NLRIs per RFC 7606. Operators SHOULD restrict SR Policy SAFI sessions to trusted controllers to limit the blast radius of misconfiguration.

9.3. Convergence and Reconvergence

When a constituent SR Policy transitions from valid to invalid (e.g., due to a link failure or BFD session down), the SRPM updates the affected FC entries per Section 6.2 without waiting for a BGP update. The headend need not withdraw and re-advertise the parent Composite CP; liveness is handled locally. When the constituent recovers, the SRPM re-installs the original FC-to-color mapping.

9.4. Observability

For operational visibility into installed Composite CP state -- including constituent validity and FC-array contents -- operators are RECOMMENDED to use the BGP-LS SR Policy distribution mechanism [RFC9857] combined with a Composite CP companion as outlined in Section 8.3. The SRPM SHOULD expose Composite CP state in the operational data model (see Section 8.5 for YANG companion status).

10. Security Considerations

The security considerations of BGP [RFC4271], BGP SR Policy [RFC9830], and SR Policy Architecture [RFC9256] apply in full to this document.

Unauthorized policy injection. As with all SR Policy BGP updates, unauthorized injection of Composite CP advertisements could redirect traffic across unintended paths. Mitigations: (a) BGP session integrity via TCP-AO [RFC5925]; (b) Route Target communities to restrict SR Policy SAFI propagation to intended headends; (c) operational policies restricting SR Policy SAFI peering to trusted controllers. Note that BGPsec [RFC8205] protects AS_PATH integrity and does not protect the Tunnel Encapsulation Attribute carrying SR Policy contents; RPKI/ROV [RFC6811] validates origin AS, not policy contents. These tools are useful complements but not direct mitigations for attribute-level injection.

FC value manipulation. An adversary able to inject or modify SR Policy SAFI BGP updates could alter FC-to-color mappings, steering specific traffic classes onto unintended paths. BGP session integrity per TCP-AO [RFC5925] is the primary mitigation.

Constituent reference amplification. A malicious or misconfigured controller can advertise large numbers of Composite CPs referencing nonexistent constituents, causing SRPM state consumption and IGP-fallback installation for each invalid constituent entry. Operators MUST apply per-peer SR Policy SAFI session limits and per-peer CP count limits as discussed in Section 9.2.

Privacy. SR Policy content reveals traffic engineering intent and topology preferences. Composite CP advertisements expose QoS class mappings (FC-to-color) and the SRv6 SID structure of constituent paths. This content MUST NOT be distributed via untrusted BGP peers. Route Target communities and peer access policies SHOULD be configured to prevent leakage beyond intended headend sets.

TCP-AO [RFC5925] provides integrity and authentication for the BGP session but does not provide confidentiality; an on-path observer can still read SR Policy contents in transit. Where confidentiality of TE intent is required, operators SHOULD additionally rely on physical/link-layer protection of BGP-carrying links, or carry BGP over an authenticated and encrypted transport (e.g., IPsec) between trusted peers.

No data-plane security changes. This document does not alter SRv6 data-plane forwarding or SID processing. Data-plane security is governed by [RFC8754] and [RFC8986].

11. IANA Considerations

11.1. BGP Tunnel Encapsulation Attribute Sub-TLVs Registry

IANA is requested to allocate one value from the "BGP Tunnel Encapsulation Attribute Sub-TLVs" registry for the Constituent SR Policy Sub-TLV defined in this document:

Value	Description	Reference
TBA1	Constituent SR Policy Sub-TLV	This document

Table 3

Per the registration procedures of that registry (RFC 9012, Section 13.1; Length-field width per RFC 9012, Section 2), the registry is partitioned by Type value range, with each range carrying both a Length-field width and a registration policy:

Range	Length field	Registration policy
1-63	1 octet	Standards Action
64-125	1 octet	First Come First Served
126-127	1 octet	Experimental Use
128-191	2 octets	Standards Action
192-252	2 octets	First Come First Served
253-254	2 octets	Experimental Use
0, 255	--	Reserved

Table 4

The Constituent SR Policy sub-TLV uses a 1-octet Length field (Section 4.2). TBA1 MUST therefore be assigned from the range *1-63* (Standards Action, 1-octet Length). The draft does not prescribe a specific numeric value within that range.

11.2. Constituent SR Policy Sub-TLV Types Registry

IANA is requested to create a new sub-registry titled *"Constituent SR Policy Sub-TLV Types"* under the "Border Gateway Protocol (BGP) Tunnel Encapsulation" registry group. Registration policies:

Range	Policy
0	Reserved
1-127	Standards Action
128-254	Expert Review
255	Reserved

Table 5

Initial assignments:

Value	Description	Reference
0	Reserved	This document
TBA2	Per-Flow FC Sub-TLV	This document
Remaining values in 1-127	Unassigned (Standards Action)	
128-254	Unassigned (Expert Review)	
255	Reserved	This document

Table 6

TBA2 is to be assigned by IANA from the range 1-127 (Standards Action) of this new sub-registry. All values in 1-127 other than TBA2 remain unassigned and subject to Standards Action. The Expert Review range (128-254) allows vendors to register experimental or vendor-specific nested sub-TLVs without requiring a Standards Track document, mirroring RFC 9012 practice.

12. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [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>>.
- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <<https://www.rfc-editor.org/info/rfc5925>>.
- [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>>.
- [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>>.
- [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>>.
- [RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <<https://www.rfc-editor.org/info/rfc9256>>.
- [RFC9830] Previdi, S., Filsfils, C., Talaulikar, K., Ed., Mattes, P., and D. Jain, "Advertising Segment Routing Policies in BGP", RFC 9830, DOI 10.17487/RFC9830, September 2025, <<https://www.rfc-editor.org/info/rfc9830>>.

13. Informative References

- [I-D.ietf-pce-multipath]
Koldychev, M., "PCEP Extensions for Signaling Multipath Information", Work in Progress, Internet-Draft, draft-

ietf-pce-multipath,
<<https://datatracker.ietf.org/doc/html/draft-ietf-pce-multipath>>.

[I-D.ietf-spring-sr-policy-yang]

SPRING Working Group, "YANG Data Model for Segment Routing Policy", Work in Progress, Internet-Draft, draft-ietf-spring-sr-policy-yang,
<<https://datatracker.ietf.org/doc/html/draft-ietf-spring-sr-policy-yang>>.

[I-D.jiang-idr-sr-policy-composite-path]

Jiang, W., Lin, C., and R. Chen, "BGP Extensions of SR Policy for Composite Candidate Path", Work in Progress, Internet-Draft, draft-jiang-idr-sr-policy-composite-path,
<<https://datatracker.ietf.org/doc/html/draft-jiang-idr-sr-policy-composite-path>>.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998,
<<https://www.rfc-editor.org/info/rfc2474>>.

[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>>.

[RFC4724] Sangli, S., Chen, E., Fernando, R., Scudder, J., and Y. Rekhter, "Graceful Restart Mechanism for BGP", RFC 4724, DOI 10.17487/RFC4724, January 2007,
<<https://www.rfc-editor.org/info/rfc4724>>.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010,
<<https://www.rfc-editor.org/info/rfc5880>>.

[RFC6811] Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R. Austein, "BGP Prefix Origin Validation", RFC 6811, DOI 10.17487/RFC6811, January 2013,
<<https://www.rfc-editor.org/info/rfc6811>>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <<https://www.rfc-editor.org/info/rfc7432>>.

- [RFC8205] Lepinski, M., Ed. and K. Sriram, Ed., "BGPsec Protocol Specification", RFC 8205, DOI 10.17487/RFC8205, September 2017, <<https://www.rfc-editor.org/info/rfc8205>>.
- [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>>.
- [RFC9494] Uttaro, J., Chen, E., Decraene, B., and J. Scudder, "Long-Lived Graceful Restart for BGP", RFC 9494, DOI 10.17487/RFC9494, November 2023, <<https://www.rfc-editor.org/info/rfc9494>>.
- [RFC9857] Previdi, S., Talaulikar, K., Ed., Dong, J., Gredler, H., and J. Tantsura, "Advertisement of Segment Routing Policies Using BGP - Link State", RFC 9857, DOI 10.17487/RFC9857, October 2025, <<https://www.rfc-editor.org/info/rfc9857>>.
- [RFC9862] Koldychev, M., Sivabalan, S., Sidor, S., Barth, C., Peng, S., and H. Bidgoli, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing (SR) Policy Candidate Paths", RFC 9862, DOI 10.17487/RFC9862, October 2025, <<https://www.rfc-editor.org/info/rfc9862>>.

Authors' Addresses

Selvamani Ramasamy (editor)
Nexthop AI
United States of America
Email: selva@nexthop.ai

Manoharan Sundaramoorthy
Nexthop AI
India
Email: manoharan@nexthop.ai

Shyam Sethuram
Nexthop AI
India
Email: shyam@nexthop.ai

Venkitraman Kasiviswanathan
Nexthop AI
United States of America
Email: venkit@nexthop.ai