Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 3, 2017

S. Previdi, Ed.
C. Filsfils
A. Sreekantiah
S. Sivabalan
Cisco Systems, Inc.
P. Mattes
Microsoft
E. Rosen
Juniper Networks
S. Lin
Google

October 30, 2016

Advertising Segment Routing Traffic Engineering Policies in BGP
draft-previdi-idr-segment-routing-te-policy-02

Abstract

This document defines a new BGP SAFI with a new NLRI in order to advertise a Segment Routing Traffic Engineering Policy (SR TE Policy). An SR TE Policy is a set of explicit paths represented by one or more segment lists. The SR TE Policy is advertised along with the Tunnel Encapsulation Attribute for which this document also defines new sub-TLVs. An SR TE policy is advertised with the information that will be used by the node receiving the advertisement in order to instantiate the policy in its forwarding table and to steer traffic according to the policy.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Segment Routing (SR) technology leverages the source routing and tunneling paradigms. [I-D.ietf-spring-segment-routing] describes the SR architecture. [I-D.ietf-spring-segment-routing-mpls] describes its instantiation on the MPLS data plane and [I-D.ietf-6man-segment-routing-header] describes the Segment Routing instantiation over the IPv6 data plane.

This document defines the Segment Routing Traffic Engineering Policy (SR TE Policy) as a set of unequal equal cost multi-path (UCMP) segment lists (representing explicit paths) as well as the mechanism allowing a router to steer traffic into an SR TE Policy.

The SR TE Policy is advertised in the Border Gateway Protocol (BGP) by the BGP speaker being a router or a controller and using extensions defined in this document. Among the information encoded in the BGP message and representing the SR TE Policy, the steering mechanism makes also use of the Extended Color Community currently defined in [I-D.ietf-idr-tunnel-encaps]

Typically, a controller defines the set of policies and advertise them to BGP 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 the receiver. In other words, a policy advertised to a given BGP speaker has significance only for that particular router and is not intended to be propagated anywhere else. Then, the receiver of the policy instantiate the policy in its routing and forwarding tables and steer traffic into it based on both the policy and destination prefix color and next-hop.

Alternatively, a router (i.e.: an BGP egress router) advertises SR TE Policies representing paths to itself. These advertisements are sent to BGP ingress nodes who instantiate these policies and steer traffic into them according to the color and endpoint/BGP next-hop of both the policy and the destination prefix.

An SR TE Policy being intended only for the receiver of the advertisement, the SR TE Policies are sent directly to each receiver and, in most of the cases will not traverse any Route Reflector (RR, [RFC4456]).

However, there are cases where a SR TE Policy is intended to a group of nodes. Also, in a deployment scenario, a controller may also rely on the standard BGP update propagation scheme which makes use of route reflectors. This cases require mechanisms that:

- o Uniquely identify each instance of a given policy.

- o Uniquely identify the intended receiver of a given SR TE Policy advertisement.

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

- o A new Subsequent Address Family Identifier (SAFI) identifying the content of the BGP message (i.e.: the SR TE Policy).

- o A new NLRI identifying the SR TE Policy.

- o A set of new TLVs to be inserted into the Tunnel Encapsulation Attribute (as defined in [I-D.ietf-idr-tunnel-encaps]) and describing the SR TE Policy.

- o An IPv4 address format route-target extended community ([RFC4360]) attached to the SR TE Policy advertisement and that indicates the intended receiver of such SR TE Policy advertisement.

- o The Extended Color Community (as defined in [I-D.ietf-idr-tunnel-encaps]) and used in order to steer traffic into an SR TE Policy.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. SR TE Policy Encoding

2.1. SR TE Policy SAFI and NLRI

A new SAFI is defined: the SR TE Policy SAFI (codepoint suggested value 73, to be assigned by IANA).

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

```
+------------------------------------------------+
|           Distinguisher (4 octets)             |
+------------------------------------------------+
|           Policy Color (4 octets)              |
+------------------------------------------------+
|           Endpoint (4 or 16 octets)            |
+------------------------------------------------+
```

where:

- Distinguisher: 4-octet value uniquely identifying the policy in the context of <color, endpoint> tuple. The distinguisher has no semantic and it's solely used by the SR TE Policy originator in order to make unique (from a NLRI perspective) multiple occurrences of the same SR TE Policy.

- Policy Color: 4-octet value identifying (with the endpoint) the policy. The color is used to match the color of the destination prefixes in order to steer traffic into the SR TE Policy.

- Endpoint: identifies the endpoint of a policy. The Endpoint may represent a single node or a set of nodes (e.g.: an anycast address or a summary address). The Endpoint is an IPv4 (4-octet) address or an IPv6 (16-octet) address according to the AFI of the NLRI.

The NLRI containing the SR TE Policy is carried in a BGP UPDATE message [RFC4271] using BGP multiprotocol extensions [RFC4760] with an AFI of 1 or 2 (IPv4 or IPv6) and with a SAFI of 73 (suggested value, to be assigned by IANA).

An update message that carries the MP_REACH_NLRI or MP_UNREACH_NLRI attribute with the SR TE 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 of the SR TE Policy SAFI NLRI is set based on the AFI. For example, if the AFI is set to IPv4 (1), then the next-hop is encoded as a 4-byte IPv4 address. If the AFI is set to IPv6 (2), then the next-hop is encoded as a 16-byte IPv6 address of the router. It is important to note that any BGP speaker receiving a BGP message with an SR TE Policy NLRI, will process it only if the NLRI is a best path as per the BGP best path selection algorithm.

2.2. SR TE Policy and Tunnel Encapsulation Attribute

The content of the SR TE Policy is encoded in the Tunnel Encapsulation Attribute originally defined in [I-D.ietf-idr-tunnel-encaps] using a new Tunnel-Type TLV (suggested codepoint is 14, to be assigned by IANA).

The SR TE Policy Encoding structure is as follows:

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

where:

- o SR TE 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 a suggested value of 14 (to be assigned by IANA).

- o Preference, Binding SID, Weight, Segment and Segment-List are new sub-TLVs defined in this document.

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

A single occurrence of "Tunnel Type: SR TE Policy" MUST be encoded within the same Tunnel Encapsulation Attribute.

Multiple occurrences of "Segment List" MAY be encoded within the same SR TE Policy.

Multiple occurrences of "Segment" MAY be encoded within the same Segment List.

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 TE Policy encodings.

If present, the Remote Endpoint sub-TLV MUST match the Endpoint of the SR TE Policy SAFI NLRI.

If present, the Color sub-TLV MUST match the Policy Color of the SR TE Policy SAFI NLRI.

2.4. SR TE Policy Sub-TLVs

This section defines the SR TE Policy sub-TLVs.

2.4.1. Preference sub-TLV

The Preference sub-TLV is used in order to determine the preference among multiple SR TE Policy originators.

The Preference sub-TLV is optional, MAY appear only once in the SR TE Policy and has following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Preference (4 octets)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- Type: to be assigned by IANA (suggested value is 6).
- Length: 6.
- Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.
- RESERVED: 1 octet of reserved bits. SHOULD be unset on transmission and MUST be ignored on receipt.
- Preference: a 4-octet value. The highest value is preferred.

The Preference is used when the same <color,endpoint> policy is advertised by multiple originators of the same SR TE Policy. The

Preference is used by the receiver in order to determine which of the received policies are to be installed. The following rules apply to the Preference:

- Preference is to be applied to the <color,endpoint> tuple. The Distinguisher MUST NOT be considered.

- Preference is used in order to determine which instance of a given SR TE Policy is to be installed. However, Preference MUST NOT influence the BGP selection algorithm and propagation rules. In other words, the preference selection happens after the BGP path selection.

2.4.2. SR TE Binding SID Sub-TLV

The Binding SID sub-TLV requests the allocation of a Binding Segment identifier associated with the SR TE Policy.

The Binding SID sub-TLV is optional, MAY appear only once in the SR TE Policy and has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Binding SID (variable, optional)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- Type: to be assigned by IANA (suggested value is 7).

- Length: specifies the length of the value field not including Type and Length fields. Can be 2 or 6 or 18.

- Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.

- RESERVED: 1 octet of reserved bits. SHOULD be unset on transmission and MUST be ignored on receipt.

- Binding SID: if length is 2, then no Binding SID is present. If length is 6 then the Binding SID contains a 4-octet SID. If length is 18 then the Binding SID contains a 16-octet IPv6 SID.

The Binding SID sub-TLV is used to instruct the receiver of the BGP message to allocate a Binding SID to the SR TE Policy. The

allocation of the Binding SID in the receiver is done according to following rules:

- o If length is 2 (no value field is present), then the receiver MUST allocate a local Binding SID whose value is chosen by the receiver.

- o If length is 6, then the value field contains the 4-octet Binding SID value the receiver SHOULD allocate.

- o If length is 18, then the value field contains the 16-octet Binding SID value the receiver SHOULD allocate.

When a controller is used in order to define and advertise SR TE Policies and when the Binding SID is allocated by the receiver, such Binding SID SHOULD be reported to the controller. The mechanisms and/or APIs used for the reporting of the Binding SID are outside the scope of this document.

Further use of the Binding SID is described in a subsequent section.

2.4.3. Weight Sub-TLV

The Weight sub-TLV specifies the weight associated to a given path (i.e.: a given segment list). The weight is used in order to apply weighted UCMP mechanism when steering traffic into a policy that includes multiple Segment Lists sub-TLVs (i.e.: multiple explicit paths).

The Weight sub-TLV is optional, MAY only appear once in the Segment List sub-TLV, and has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              Weight                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: to be assigned by IANA (suggested value is 9).

Length: 6.

Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.

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

When present, the Weight sub-TLV specifies a weight to be associated with the corresponding Segment List, for use in unequal cost multi-path. Weights are applied by summing the total value of all of the weights for all Segment Lists, and then assigning a fraction of the forwarded traffic to each Segment List in proportion its weight's fraction of the total.

2.4.4. Segment List Sub-TLV

The Segment List sub-TLV is used in order to encode a single explicit path towards the endpoint. The Segment List sub-TLV includes the elements of the paths (i.e.: segments) as well as an optional Weight 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 code point defines the size of the length field. Therefore, for the Segment List sub-TLV a code point of 128 (suggested value, to be assigned by IANA) is used.

The Segment List sub-TLV is mandatory, MAY appear multiple times in the SR TE Policy and has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |             Length            |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                           sub-TLVs                          //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- o Type: to be assigned by IANA (suggested value is 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 unset on transmission and MUST be ignored on receipt.
- o sub-TLVs:
  - * An optional single Weight sub-TLV.

* One or more Segment sub-TLVs.

The Segment List sub-TLV is mandatory.

Multiple occurrences of the Segment List sub-TLV MAY appear in the SR TE Policy.

When multiple occurrences of the Segment List sub-TLV appear in the SR TE Policy, the traffic is load-balanced across them either through an ECMP scheme (if no Weight sub-TLV is present) or through a weighted UCMP scheme according to Section 2.4.3.

2.4.5. Segment Sub-TLV

The Segment sub-TLV describes a single segment in a segment list (i.e.: a single element of the explicit path). Multiple Segment sub-TLVs constitute an explicit path of the SR TE Policy.

The Segment sub-TLV is mandatory and MAY appear multiple times in the Segment List sub-TLV.

This document defines 8 different types of Segment Sub-TLVs:

Type 1: SID only, in the form of MPLS Label
Type 2: SID only, in the form of IPv6 address
Type 3: IPv4 Node Address with optional SID
Type 4: IPv6 Node Address with optional SID
Type 5: IPv4 Address + index with optional SID
Type 6: IPv4 Local and Remote addresses with optional SID
Type 7: IPv6 Address + index with optional SID
Type 8: IPv6 Local and Remote addresses with optional SID

2.4.5.1. Type 1: SID only, in the form of MPLS Label

The Type-1 Segment Sub-TLV encodes a single SID in the form of an MPLS label. The format is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Label                        | TC  |S|       TTL     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

o Type: suggested value 1, to be assigned by IANA.

- Length is 6.

- Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.

- RESERVED: 1 octet of reserved bits. SHOULD be unset on transmission and MUST be ignored on receipt.

- Label: 20 bits of label value.

- TC: 3 bits of traffic class.

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

- TTL: 1 octet of TTL.

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

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

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

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

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

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

2.4.5.2. Type 2: SID only, in the form of IPv6 address

The Type-2 Segment Sub-TLV encodes a single SID in the form of an IPv6 address. The format is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                       IPv6 SID (16 octets)                  //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- o Type: suggested value 2, to be assigned by IANA.
- o Length is 18.
- o Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.
- o RESERVED: 1 octet of reserved bits. SHOULD be unset on transmission and MUST be ignored on receipt.
- o IPv6 SID: 16 octets of IPv6 address.

The IPv6 Segment Identifier (IPv6 SID) is defined in [I-D.ietf-6man-segment-routing-header].

2.4.5.3. Type 3: IPv4 Node Address with optional SID

The Type-3 Segment Sub-TLV encodes an IPv4 node address and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 IPv4 Node Address (4 octets)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                SID (optional, 4 or 16 octets)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- o Type: suggested value 3, to be assigned by IANA.
- o Length is 6 or 10 or 22.

- Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.

- RESERVED: 1 octet of reserved bits. SHOULD be unset on transmission and MUST be ignored on receipt.

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

- SID: either 4 octet MPLS SID or a 16 octet IPv6 address.

The following applies to the Type-3 Segment sub-TLV:

- The IPv4 Node Address MUST be present.

- The SID is optional and MAY be of one of the following formats:

  * MPLS SID: a 4 octet label containing label, TC, S and TTL as defined in Section 2.4.5.1.

  * IPV6 SID: a 16 octet IPv6 address.

- If length is 6, then only the IPv4 Node Address is present.

- If length is 10, then the IPv4 Node Address and the MPLS SID are present.

- If length is 22, then the IPv4 Node Address and the IPv6 SID are present.

2.4.5.4. Type 4: IPv6 Node Address with optional SID

The Type-4 Segment Sub-TLV encodes an IPv6 node address and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                IPv6 Node Address (16 octets)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                SID (optional, 4 or 16 octets)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- Type: suggested value 4, to be assigned by IANA.

- o Length is 18 or 22 or 34.

- o Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.

- o RESERVED: 1 octet of reserved bits. SHOULD be unset on transmission and MUST be ignored on receipt.

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

- o SID: either 4 octet MPLS SID or a 16 octet IPv6 address.

The following applies to the Type-4 Segment sub-TLV:

- o The IPv6 Node Address MUST be present.

- o The SID is optional and MAY be of one of the following formats:

  * MPLS SID: a 4 octet label containing label, TC, S and TTL as defined in Section 2.4.5.1.

  * IPV6 SID: a 16 octet IPv6 address.

- o If length is 18, then only the IPv6 Node Address is present.

- o If length is 22, then the IPv6 Node Address and the MPLS SID are present.

- o If length is 34, then the IPv6 Node Address and the IPv6 SID are present.

2.4.5.5. Type 5: IPv4 Address + index with optional SID

The Type-5 Segment Sub-TLV encodes an IPv4 node address, an interface index (IfIndex) and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 IfIndex (4 octets)                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 IPv4 Node Address (4 octets)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                SID (optional, 4 or 16 octets)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- o Type: suggested value 5, to be assigned by IANA.
- o Length is 10 or 14 or 26.
- o Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.
- o RESERVED: 1 octet of reserved bits. SHOULD be unset on transmission and MUST be ignored on receipt.
- o IfIndex: 4 octets of interface index.
- o IPv4 Node Address: a 4 octet IPv4 address representing a node.
- o SID: either 4 octet MPLS SID or a 16 octet IPv6 address.

The following applies to the Type-5 Segment sub-TLV:

- o The IPv4 Node Address MUST be present.
- o The Interface Index (IfIndex) MUST be present.
- o The SID is optional and MAY be of one of the following formats:
    - * MPLS SID: a 4 octet label containing label, TC, S and TTL as defined in Section 2.4.5.1.
    - * IPV6 SID: a 16 octet IPv6 address.
- o If length is 10, then the IPv4 Node Address and IfIndex are present.
- o If length is 14, then the IPv4 Node Address, the IfIndex and the MPLS SID are present.

o If length is 26, then the IPv4 Node Address, the IfIndex and the IPv6 SID are present.

2.4.5.6. Type 6: IPv4 Local and Remote addresses with optional SID

The Type-6 Segment Sub-TLV encodes an IPv4 node address, an adjacency local address, an adjacency remote address and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Local IPv4 Address (4 octets)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Remote IPv4 Address  (4 octets)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                     SID (4 or 16 octets)                    //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

o Type: suggested value 6, to be assigned by IANA.

o Length is 10 or 14 or 26.

o Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.

o RESERVED: 1 octet of reserved bits. SHOULD be unset 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 SID: either 4 octet MPLS SID or a 16 octet IPv6 address.

The following applies to the Type-6 Segment sub-TLV:

o The Local IPv4 Address MUST be present and represents an adjacency local address.

o The Remote IPv4 Address MUST be present and represents the remote end of the adjacency.

o The SID is optional and MAY be of one of the following formats:

   * MPLS SID: a 4 octet label containing label, TC, S and TTL as defined in Section 2.4.5.1.

   * IPV6 SID: a 16 octet IPv6 address.

o If length is 10, then only the IPv4 Local and Remote addresses are present.

o If length is 14, then the IPv4 Local address, IPv4 Remote address and the MPLS SID are present.

o If length is 26, then the IPv4 Local address, IPv4 Remote address and the IPv6 SID are present.

2.4.5.7. Type 7: IPv6 Address + index with optional SID

The Type-7 Segment Sub-TLV encodes an IPv6 node address, an interface index and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 IfIndex (4 octets)                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                IPv6 Node Address (16 octets)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                SID (optional, 4 or 16 octets)               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

o Type: suggested value 7, to be assigned by IANA.

o Length is 22 or 26 or 38.

o Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.

o RESERVED: 1 octet of reserved bits. SHOULD be unset on transmission and MUST be ignored on receipt.

o IfIndex: 4 octets of interface index.

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

- o SID: either 4 octet MPLS SID or a 16 octet IPv6 address.

The following applies to the Type-7 Segment sub-TLV:

- o The IPv6 Node Address MUST be present.

- o The Interface Index MUST be present.

- o The SID is optional and MAY be of one of the following formats:

  * MPLS SID: a 4 octet label containing label, TC, S and TTL as defined in Section 2.4.5.1.

  * IPV6 SID: a 16 octet IPv6 address.

- o If length is 22, then the IPv6 Node Address and IfIndex are present.

- o If length is 26, then the IPv6 Node Address, the IfIndex and the MPLS SID are present.

- o If length is 38, then the IPv6 Node Address, the IfIndex and the IPv6 SID are present.

2.4.5.8. Type 8: IPv6 Local and Remote addresses with optional SID

The Type-8 Segment Sub-TLV encodes an IPv6 node address, an adjacency local address, an adjacency remote address and an optional SID in the form of either an MPLS label or an IPv6 address. The format is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |   Length      |     Flags     |   RESERVED    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//               Local IPv6 Address (16 octets)                //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//               Remote IPv6 Address  (16 octets)              //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                     SID (4 or 16 octets)                    //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- o Type: suggested value 8, to be assigned by IANA.

- o Length is 34 or 38 or 50.

- o Flags: 1 octet of flags. None is defined at this stage. Flags SHOULD be unset on transmission and MUST be ignored on receipt.

- o RESERVED: 1 octet of reserved bits. SHOULD be unset 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 SID: either 4 octet MPLS SID or a 16 octet IPv6 address.

The following applies to the Type-8 Segment sub-TLV:

- o The Local IPv6 Address MUST be present and represents an adjacency local address.

- o The Remote IPv6 Address MUST be present and represents the remote end of the adjacency.

- o The SID is optional and MAY be of one of the following formats:

  * MPLS SID: a 4 octet label containing label, TC, S and TTL as defined in Section 2.4.5.1.

  * IPV6 SID: a 16 octet IPv6 address.

- o If length is 34, then only the IPv6 Local and Remote addresses are present.

- o If length is 38, then the IPv6 Local address, IPv4 Remote address and the MPLS SID are present.

- o If length is 50, then the IPv6 Local address, IPv4 Remote address and the IPv6 SID are present.

3. SR TE Policy Operations

3.1. Configuration and Advertisement of SR TE Policies

Typically, but not limited to, a SR TE Policy is configured into a controller and on the base of each receiver. In other words, each SR TE Policy configured is related to the intended receiver. It is therefore normal for a given <color,endpoint> SR TE Policy to have

multiple instances with different content (i.e.: different segment lists) where each of these instances (of the same policy) is intended to be sent to different receivers.

Each instance of the same SR TE Policy will have a different Distinguisher in order to prevent BGP selection among these instances along the distribution of BGP updates.

Moreover, a Route-Target extended community SHOULD be attached to the SR TE Policy and that identifies the intended receiver of the advertisement.

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

3.2. Multipath Operation

The SR TE Policy MAY contain multiple Segment Lists which, in the absence of the Weight TLV, signifies equal cost load balancing amongst them.

When a weight sub-TLV is encoded in each Segment List TLV, then the weight value SHOULD be used in order to perform an unequal cost load balance amongst the Segment Lists as specified in Section 2.4.3.

3.3. Binding SID TLV

When the optional Binding SID sub-TLV is present, it indicates an instruction, to the receiving BGP speaker to allocate a Binding SID for the list of SIDs the Binding sub-TLV is related to.

Any incoming packet with the Binding SID as active segment (according to the terminology described in [I-D.ietf-spring-segment-routing]) will then have the Binding SID swapped with the list of SIDs specified in the Segment List sub-TLVs on the allocating BGP speaker. The allocated Binding SID MAY be then advertised by the BGP speaker that created it, through, e.g., BGP-LS in order to, typically, feed a controller with the updated topology and SR TE Policy information.

3.4. Reception of an SR TE Policy

On reception of a SR TE Policy, a BGP speaker MUST determine if the SR TE Policy is first acceptable, then usable.

While only usable SR TE Policies are instantiated, acceptable SR TE Policies (i.e.: also the non-usable ones) MAY be propagated.

Any SR TE Policy update that has been determined acceptable is kept in the BGP database. This includes non-usable SR TE Policies.

3.4.1. Acceptance of a SR TE Policy Update

When a BGP speaker receives an SR TE Policy from a neighbor it has to determine if the SR TE Policy advertisement is acceptable. The following applies:

- o The SR TE Policy NLRI MUST have a color value and MUST have an endpoint value.

- o The SR TE Policy NLRI MUST have distinguisher field.

- o The SR TE Policy update MUST have either the NO_ADV community or at least one route-target extended community in IPv4-address format.

- o The Tunnel Encapsulation Attribute MUST be attached to the BGP Update and MUST have the Tunnel Type set to SR TE Policy (value to be assigned by IANA).

- o Within the SR TE Policy, at least one Segment List sub-TLV MUST be present.

- o Within the Segment List sub-TLV at least one Segment sub-TLV MUST be present.

The Remote Endpoint and Color sub-TLVs, as defined in [I-D.ietf-idr-tunnel-encaps], MAY also be present in the SR TE Policy encodings. If present, the Remote Endpoint sub-TLV MUST match the Endpoint of the SR TE Policy SAFI NLRI. If they don't match, the SR TE Policy advertisement MUST be considered as not acceptable. If present, the Color sub-TLV MUST match the Policy Color of the SR TE Policy SAFI NLRI. If they don't match, the SR TE Policy advertisement MUST be considered as not acceptable.

A non-acceptable SR TE Policy update that has a valid NLRI portion with invalid attribute portion MUST be considered as a withdraw of the SR TE Policy.

A non-acceptable SR TE Policy update that has an invalid NLRI portion MUST trigger a reset of the BGP session.

3.4.2. Usable SR TE Policy

When the receiver has determined that the received SR TE Policy is acceptable according to previous section, it has to determine if the received SR TE Policy is usable.

The receiver MUST check whether route-target or NO_ADVERTISE communities are attached to it. If no route-target is present and the NO_ADVERTISE community is present, then the SR TE Policy is usable.

If one or more route-targets are present, then at least one route-target MUST match the BGP Identifier (BGP Router-ID) of the receiver in order 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 no one matches the local BGP router-ID, then, while the SR TE Policy is acceptable, the SR TE Policy is not usable. It has to be noted that if the receiver has been explicitly configured to do so, it MAY propagate the SR TE Policy to its neighbors as defined in Section 3.4.4.

The following applies to usable SR TE Policies:

- Any segment sub-TLV of type 3 to 8 that is present in the segment list MUST be either validated or resolved:

  if the SID portion of the sub-TLV is present, then the segment MUST be validated by the receiver. Validation consists of verifying that the SID value is related to the network address.

  if the SID portion of the sub-TLV is not present, then the segment MUST be resolved by the receiver. Resolution consists of taking from the receiver database (e.g.; from the link-state or routing information base) that the SID value related to the network address in the sub-TLV.

- The receiver MUST check the validity of the first SID of each Segment List sub-TLV of the SR TE Policy. The first SID MUST be known in the receiver local table either as a label (in the case the SID encodes a label value) or as an IPv6 address.

- Any invalid segment of the segment list MUST cause an invalidation of the whole segment list. However, the SR TE Policy is still usable if at least one segment list is valid.

- The receiver much keep track of the validated segment (i.e.: the first segment and any segment encoded in Sub-TLV type 3 to 8). A segment who failed validation may become valid after a network event and vice versa. The receiver SHOULD keep the state of the received SR TE Policies based on latest state of the segments requires validation.

It has to be noted that an SR-TE policy may be received by a server that is not a router, and that does not have the necessary state that allows him to infer the next-hop from the first segment. In that case, if the server needs to send a packet according to a particular SR-TE policy, it SHOULD push on the label stack that the policy specifies, and then send the packet to a default router (or default gateway).

3.4.3. Instantiation of an SR TE Policy

On reception of an acceptable, valid and usable SR TE Policy, a BGP speaker SHOULD instantiate the SR TE Policy in its routing and forwarding table with the set of segment lists (i.e.: explicit paths) included in the policy and taking into account the Binding SID and Weight sub-TLVs.

The receiver of the SR TE Policy SHOULD program its MPLS or IPv6 data planes so that BGP destination prefixes matching their Extended Color Community and BGP next-hop with the SR TE Policy SAFI NLRI Color and Endpoint are steered into the SR TE Policy and forwarded accordingly.

When building the MPLS label stack or the IPv6 Segment list from the Segment List sub-TLV, the receiving BGP speaker MUST interpret the set of Segment sub-TLVs as follows:

- The first Segment sub-TLV represents the topmost label or the first IPv6 segment. In the receiving BGP speaker, it identifies the first segment the traffic will be directed towards to (along the SR TE explicit path).

- The last Segment sub-TLV represents the bottommost label or the last IPv6 segment.

As described in Section 2.4.3, when present, the Weight sub-TLV specifies a weight to be associated with the corresponding Segment List, for use in unequal-cost multi path. Weights are applied by summing the total value of all of the weights for all Segment Lists, and then assigning a fraction of the forwarded traffic to each Segment List in proportion its weight's fraction of the total.

If in a SR TE Policy only some of the segment lists have a Weight Sub-TLV present, then for those who haven't any weight, a value of 1 is assumed.

3.4.4. Propagation of an SR TE Policy

By default, a BGP node receiving an SR TE Policy MUST NOT not propagate it to any eBGP neighbor.

However, a node MAY be explicitly configured in order to advertise a received SR TE Policy update to neighbors according to normal BGP rules (iBGP and eBGP propagation), e.g., in the case the node is a Route-Reflector.

SR TE Policies that have been determined acceptable and valid can be propagated, even the ones that are not usable.

Only SR TE Policies that do not have the NO_ADVERTISE community attached to them can be propagated.

3.5. Steering Traffic into a SR TE Policy

The Color field of the NLRI allows association of destination prefixes with a given SR TE Policy. The BGP speaker SHOULD then attach a Color Extended Community (as defined in [RFC5512]) to destination prefixes (e.g.: IPv4/IPv6 unicast prefixes) in order to allow the receiver of the SR TE Policy and of the destination prefix to steer traffic into the SR TE Policy if the destination prefix:

- o Has a BGP next-hop matching the SR TE Policy SAFI NLRI Endpoint and

- o Has an attached Extended Color Community with the same value as the color of the SR TE Policy NLRI Color.

On the receiving BGP speaker, all destination prefixes that share the same Extended Color Community value and the same BGP next-hop are steered to the corresponding SR TE Policy that has been instantiated and which matches the Color and Endpoint NLRI values.

It is assumed that only one Extended Color Community is attached to a destination prefix. In case a destination prefix is received with multiple Extended Color Communities, the receiver MUST consider the color corresponding to the SR TE Policy having the highest Preference TLV value. In case of multiple policies having the same preference, as a breaking tie, the router SHOULD select the policy having the lowest color value.

Different destination prefixes can be steered into distinct SR TE Policies by coloring them differently.

3.6. Flowspec and SR TE Policies

The SR TE Policy can be carried in context of a Flowspec NLRI ([RFC5575]). In this case, when the redirect to IP next-hop is specified as in [I-D.ietf-idr-flowspec-redirect-ip], the tunnel to the next-hop is specified by the segment list in the Segment List sub-TLVs. The Segment List (e.g..: label stack or IPv6 segment list) is imposed to flows matching the criteria in the Flowspec route in order to steer them towards the next-hop as specified in the SR TE Policy SAFI NLRI.

4. Acknowledgments

The authors of this document would like to thank Dhanendra Jain, Shyam Sethuram, Acee Lindem and Imtiyaz Mohammad for their comments and review of this document.

5. IANA Considerations

This document defines:

- o a new SAFI in the registry "Subsequent Address Family Identifiers (SAFI) Parameters":

Suggested Value	Description	Reference
73	SR TE Policy SAFI	This document

- o a new Tunnel-Type in the registry "BGP Tunnel Encapsulation Attribute Tunnel Types":

Suggested Value	Description	Reference
14	SR TE Policy Type	This document

- o new sub-TLVs in the registry "BGP Tunnel Encapsulation Attribute sub-TLVs":

Suggested Value	Description	Reference
6	Preference sub-TLV	This document
7	Binding SID sub-TLV	This document
8	Segment List sub-TLV	This document

o A new registry called "SR TE Policy List Sub-TLVs" with following codepoints:

Suggested Value	Description	Reference
1	MPLS SID	This document
2	IPv6 SID	This document
3	IPv4 Node and SID	This document
4	IPv6 Node and SID	This document
5	IPv4 Node, index and SID	This document
6	IPv4 Local/Remote addresses and SID	This document
7	IPv6 Node, index and SID	This document
8	IPv6 Local/Remote addresses and SID	This document
9	Weight sub-TLV	This document

6. Security Considerations

TBD.

7. References

7.1. Normative References

[I-D.ietf-idr-tunnel-encaps]
Rosen, E., Patel, K., and G. Velde, "The BGP Tunnel Encapsulation Attribute", draft-ietf-idr-tunnel-encaps-02 (work in progress), May 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://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, <http://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, <http://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, <http://www.rfc-editor.org/info/rfc4760>.

[RFC5512] Mohapatra, P. and E. Rosen, "The BGP Encapsulation Subsequent Address Family Identifier (SAFI) and the BGP Tunnel Encapsulation Attribute", RFC 5512, DOI 10.17487/RFC5512, April 2009, <http://www.rfc-editor.org/info/rfc5512>.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009, <http://www.rfc-editor.org/info/rfc5575>.

7.2. Informational References

[I-D.ietf-6man-segment-routing-header] Previdi, S., Filsfils, C., Field, B., Leung, I., Linkova, J., Aries, E., Kosugi, T., Vyncke, E., and D. Lebrun, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-02 (work in progress), September 2016.

[I-D.ietf-idr-flowspec-redirect-ip] Uttaro, J., Haas, J., Texier, M., Andy, A., Ray, S., Simpson, A., and W. Henderickx, "BGP Flow-Spec Redirect to IP Action", draft-ietf-idr-flowspec-redirect-ip-02 (work in progress), February 2015.

[I-D.ietf-spring-segment-routing] Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-09 (work in progress), July 2016.

[I-D.ietf-spring-segment-routing-mpls] Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Shakir, R., jefftant@gmail.com, j., and E. Crabbe, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls-05 (work in progress), July 2016.

[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, <http://www.rfc-editor.org/info/rfc4456>.

Authors' Addresses

Stefano Previdi (editor)
Cisco Systems, Inc.
Via Del Serafico, 200
Rome 00142
Italy

Email: sprevidi@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Brussels
BE

Email: cfilsfil@cisco.com

Arjun Sreekantiah
Cisco Systems, Inc.
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: asreekan@cisco.com

Siva Sivabalan
Cisco Systems, Inc.
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: msiva@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

Steven Lin
Google

Email: stevenlin@google.com