

Internet Engineering Task Force (IETF)
Request for Comments: 8667
Category: Standards Track
ISSN: 2070-1721

S. Previdi, Ed.
Huawei Technologies
L. Ginsberg, Ed.
C. Filsfils
Cisco Systems, Inc.
A. Bashandy
Arrcus
H. Gredler
RtBrick Inc.
B. Decraene
Orange
December 2019

IS-IS Extensions for Segment Routing

Abstract

Segment Routing (SR) allows for a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called "segments". These segments are advertised by the link-state routing protocols (IS-IS and OSPF).

This document describes the IS-IS extensions that need to be introduced for Segment Routing operating on an MPLS data plane.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at <https://www.rfc-editor.org/info/rfc8667>.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction
 - 1.1. Requirements Language
2. Segment Routing Identifiers
 - 2.1. Prefix Segment Identifier (Prefix-SID) Sub-TLV
 - 2.1.1. Flags
 - 2.1.2. Prefix-SID Propagation
 - 2.2. Adjacency Segment Identifier
 - 2.2.1. Adjacency Segment Identifier (Adj-SID) Sub-TLV
 - 2.2.2. Adjacency Segment Identifier (LAN-Adj-SID) Sub-TLV
 - 2.3. SID/Label Sub-TLV

- 2.4. SID/Label Binding TLV
 - 2.4.1. Flags
 - 2.4.2. Range
 - 2.4.3. Prefix Length, Prefix
 - 2.4.4. Mapping Server Prefix-SID
 - 2.4.5. SID/Label Sub-TLV
 - 2.4.6. Example Encodings
 - 2.5. Multi-Topology SID/Label Binding TLV
 - 3. Router Capabilities
 - 3.1. SR-Capabilities Sub-TLV
 - 3.2. SR-Algorithm Sub-TLV
 - 3.3. SR Local Block Sub-TLV
 - 3.4. SRMS Preference Sub-TLV
 - 4. IANA Considerations
 - 4.1. Sub-TLVs for Types 22, 23, 25, 141, 222, and 223
 - 4.2. Sub-TLVs for Types 135, 235, 236, and 237
 - 4.3. Sub-TLVs for Type 242
 - 4.4. New TLV Codepoint and Sub-TLV Registry
 - 5. Security Considerations
 - 6. References
 - 6.1. Normative References
 - 6.2. Informative References
- Acknowledgements
Contributors
Authors' Addresses

1. Introduction

Segment Routing (SR) allows for a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called "segments". These segments are advertised by the link-state routing protocols (IS-IS and OSPF). Prefix segments represent an ECMP-aware shortest path to a prefix (or a node), as per the state of the IGP topology. Adjacency segments represent a hop over a specific adjacency between two nodes in the IGP. A prefix segment is typically a multi-hop path while an adjacency segment, in most of the cases, is a one-hop path. SR's control plane can be applied to both IPv6 and MPLS data planes and does not require any additional signaling (other than the regular IGP). For example, when used in MPLS networks, SR paths do not require any LDP or RSVP-TE signaling. Still, SR can interoperate in the presence of Label Switched Paths (LSPs) established with RSVP or LDP.

There are additional segment types, e.g., the Binding SID as defined in [RFC8402]. This document also defines an advertisement for one type of Binding SID: the Mirror Context segment.

This document describes the IS-IS extensions that need to be introduced for Segment Routing operating on an MPLS data plane.

The Segment Routing architecture is described in [RFC8402]. Segment Routing use cases are described in [RFC7855].

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Segment Routing Identifiers

The Segment Routing architecture [RFC8402] defines different types of Segment Identifiers (SIDs). This document defines the IS-IS encodings for the IGP-Prefix Segment, the IGP-Adjacency Segment, the IGP-LAN-Adjacency Segment, and the Binding Segment.

2.1. Prefix Segment Identifier (Prefix-SID) Sub-TLV

A new IS-IS sub-TLV is defined: the Prefix Segment Identifier (Prefix-SID) sub-TLV.

The Prefix-SID sub-TLV carries the Segment Routing IGP-Prefix-SID as defined in [RFC8402]. The 'Prefix-SID' MUST be unique within a given IGP domain (when the L-Flag is not set).

A Prefix-SID sub-TLV is associated to a prefix advertised by a node and MAY be present in any of the following TLVs:

TLV-135 (Extended IPv4 reachability) defined in [RFC5305].

TLV-235 (Multi-topology IPv4 Reachability) defined in [RFC5120].

TLV-236 (IPv6 IP Reachability) defined in [RFC5308].

TLV-237 (Multi-topology IPv6 IP Reachability) defined in [RFC5120].

The Binding TLV and Multi-Topology Binding TLV are defined in Sections 2.4 and 2.5, respectively.

The Prefix-SID sub-TLV has the following format:

```

0          1          2          3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
|  Type          |  Length        |  Flags         |  Algorithm     |
+-----+-----+-----+-----+
|                               SID/Index/Label (variable) |
+-----+-----+-----+-----+
```

where:

Type: 3

Length: 5 or 6 depending on the size of the SID (described below)

Flags: 1-octet field of the following flags:

```

0 1 2 3 4 5 6 7
+-----+
|R|N|P|E|V|L|
+-----+
```

where:

R-Flag: Re-advertisement Flag. If set, then the prefix to which this Prefix-SID is attached has been propagated by the router from either another level (i.e., from Level-1 to Level-2 or the opposite) or redistribution (e.g., from another protocol).

N-Flag: Node-SID Flag. If set, then the Prefix-SID refers to the router identified by the prefix. Typically, the N-Flag is set on Prefix-SIDs that are attached to a router loopback address. The N-Flag is set when the Prefix-SID is a Node-SID as described in [RFC8402].

P-Flag: No-PHP (No Penultimate Hop-Popping) Flag. If set, then the penultimate hop MUST NOT pop the Prefix-SID before delivering the packet to the node that advertised the Prefix-SID.

E-Flag: Explicit NULL Flag. If set, any upstream neighbor of the Prefix-SID originator MUST replace the Prefix-SID with a Prefix-SID that has an Explicit NULL value (0 for IPv4 and 2 for IPv6) before forwarding the packet.

V-Flag: Value Flag. If set, then the Prefix-SID carries a value (instead of an index). By default, the flag is UNSET.

L-Flag: Local Flag. If set, then the value/index carried by the Prefix-SID has local significance. By default, the flag is UNSET.

Other bits: MUST be zero when originated and ignored when received.

Algorithm: the router may use various algorithms when calculating reachability to other nodes or to prefixes attached to these nodes. Algorithm identifiers are defined in Section 3.2. Examples of these algorithms are metric-based Shortest Path First (SPF), various sorts of Constrained SPF, etc. The Algorithm field of the Prefix-SID contains the identifier of the algorithm the router uses to compute the reachability of the prefix to which the Prefix-SID is associated.

At origination, the Prefix-SID Algorithm field MUST be set to 0 or to any value advertised in the SR-Algorithm sub-TLV (see Section 3.2).

A router receiving a Prefix-SID from a remote node and with an algorithm value that such remote node has not advertised in the SR-Algorithm sub-TLV (see Section 3.2) MUST ignore the Prefix-SID sub-TLV.

SID/Index/Label as defined in Section 2.1.1.1.

When the Prefix-SID is an index (and the V-Flag is not set), the value is used to determine the actual label value inside the set of all advertised label ranges of a given router. This allows a receiving router to construct the forwarding state to a particular destination router.

In many use cases, a 'stable transport' address is overloaded as an identifier of a given node. Because Prefixes may be re-advertised into other levels, there may be some ambiguity (e.g., originating router vs. L1L2 router) for which node a particular IP prefix serves as the identifier. The Prefix-SID sub-TLV contains the necessary flags to disambiguate Prefix-to-node mappings. Furthermore, if a given node has several 'stable transport' addresses, there are flags to differentiate those among other Prefixes advertised from a given node.

2.1.1.1. Flags

2.1.1.1.1. V-Flag and L-Flag

The V-Flag indicates whether the SID/Index/Label field is a value or an index.

The L-Flag indicates whether the value/index in the SID/Index/Label field has local or global significance.

The following settings for V-Flag and L-Flag are valid:

The V-Flag and L-Flag are set to 0: The SID/Index/Label field is a 4-octet index defining the offset in the SID/Label space advertised by this router using the encodings defined in Section 3.1.

The V-Flag and L-Flag are set to 1: The SID/Index/Label field is a 3-octet local label where the 20 rightmost bits are used for encoding the label value.

All other combinations of V-Flag and L-Flag are invalid, and any SID advertisement received with an invalid setting for the V-Flag and

L-Flag MUST be ignored.

2.1.1.2. R-Flag and N-Flag

The R-Flag MUST be set for prefixes that are not local to the router and are advertised because of:

propagation (Level-1 into Level-2);
leaking (Level-2 into Level-1); or
redistribution (e.g., from another protocol).

In the case where a Level-1-2 router has local interface addresses configured in one level, it may also propagate these addresses into the other level. In such case, the Level-1-2 router MUST NOT set the R bit.

The N-Flag is used in order to define a Node-SID. A router MAY set the N-Flag only if all of the following conditions are met:

The prefix to which the Prefix-SID is attached is local to the router (i.e., the prefix is configured on one of the local interfaces, e.g., a 'stable transport' loopback).

The prefix to which the Prefix-SID is attached has a Prefix length of either /32 (IPv4) or /128 (IPv6).

The router MUST ignore the N-Flag on a received Prefix-SID if the prefix has a Prefix length different than /32 (IPv4) or /128 (IPv6).

The Prefix Attribute Flags sub-TLV [RFC7794] also defines the N-Flag and R-Flag and with the same semantics of the equivalent flags defined in this document. Whenever the Prefix Attribute Flags sub-TLV is present for a given prefix, the values of the N-Flag and R-Flag advertised in that sub-TLV MUST be used, and the values in a corresponding Prefix-SID sub-TLV (if present) MUST be ignored.

2.1.1.3. E-Flag and P-Flag

The following behavior is associated with the settings of the E-Flag and P-Flag:

- * If the P-Flag is not set, then any upstream neighbor of the Prefix-SID originator MUST pop the Prefix-SID. This is equivalent to the "penultimate hop-popping" mechanism used in the MPLS data plane, which improves performance of the ultimate hop. MPLS EXP bits of the Prefix-SID are not preserved to the ultimate hop (the Prefix-SID being removed). If the P-Flag is unset, the received E-Flag is ignored.
- * If the P-Flag is set, then:
 - If the E-Flag is not set, then any upstream neighbor of the Prefix-SID originator MUST keep the Prefix-SID on top of the stack. This is useful when, e.g., the originator of the Prefix-SID must stitch the incoming packet into a continuing MPLS LSP to the final destination. This could occur at an inter-area border router (prefix propagation from one area to another) or at an interdomain border router (prefix propagation from one domain to another).
 - If the E-Flag is set, then any upstream neighbor of the Prefix-SID originator MUST replace the Prefix-SID with a Prefix-SID having an Explicit NULL value. This is useful, e.g., when the originator of the Prefix-SID is the final destination for the related prefix and the originator wishes to receive the packet with the original EXP bits.

When propagating (from either Level-1 to Level-2 or Level-2 to Level-1) a reachability advertisement originated by another IS-IS speaker,

the router MUST set the P-Flag and MUST clear the E-Flag of the related Prefix-SIDs.

2.1.2. Prefix-SID Propagation

The Prefix-SID sub-TLV MUST be included when the associated Prefix Reachability TLV is propagated across level boundaries.

The Level-1-2 router that propagates the Prefix-SID sub-TLV between levels maintains the content (flags and SID), except as noted in Sections 2.1.1.2 and 2.1.1.3.

2.2. Adjacency Segment Identifier

A new IS-IS sub-TLV is defined: the Adjacency Segment Identifier (Adj-SID) sub-TLV.

The Adj-SID sub-TLV is an optional sub-TLV carrying the Segment Routing IGP-Adjacency-SID as defined in [RFC8402] with flags and fields that may be used, in future extensions of Segment Routing, for carrying other types of SIDs.

IS-IS adjacencies are advertised using one of the IS Neighbor TLVs below:

TLV-22 (Extended IS reachability) [RFC5305]

TLV-222 (MT-ISN) [RFC5120]

TLV-23 (IS Neighbor Attribute) [RFC5311]

TLV-223 (MT IS Neighbor Attribute) [RFC5311]

TLV-141 (inter-AS reachability information) [RFC5316]

Multiple Adj-SID sub-TLVs MAY be associated with a single IS Neighbor.

2.2.1. Adjacency Segment Identifier (Adj-SID) Sub-TLV

The following format is defined for the Adj-SID sub-TLV:

```

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      |      Weight      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     SID/Label/Index (variable) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

where:

Type: 31

Length: 5 or 6 depending on size of the SID

Flags: 1-octet field of the following flags:

```

0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+---+
|F|B|V|L|S|P|  |
+---+---+---+---+---+---+---+---+
```

where:

F-Flag: Address-Family Flag. If unset, then the Adj-SID is used when forwarding IPv4-encapsulated traffic to the neighbor. If set, then the Adj-SID is used when forwarding IPv6-encapsulated traffic to the neighbor.

B-Flag: Backup Flag. If set, the Adj-SID is eligible for protection (e.g., using IP Fast Reroute (IPFRR) or MPLS Fast Reroute (MPLS-FRR)) as described in [RFC8402].

V-Flag: Value Flag. If set, then the Adj-SID carries a value. By default, the flag is SET.

L-Flag: Local Flag. If set, then the value/index carried by the Adj-SID has local significance. By default, the flag is SET.

S-Flag: Set Flag. When set, the S-Flag indicates that the Adj-SID refers to a set of adjacencies (and therefore MAY be assigned to other adjacencies as well).

P-Flag: Persistent Flag. When set, the P-Flag indicates that the Adj-SID is persistently allocated, i.e., the Adj-SID value remains consistent across router restart and/or interface flap.

Other bits: MUST be zero when originated and ignored when received.

Weight: 1 octet. The value represents the weight of the Adj-SID for the purpose of load balancing. The use of the weight is defined in [RFC8402].

SID/Index/Label as defined in Section 2.1.1.1.

An SR-capable router MAY allocate an Adj-SID for each of its adjacencies.

An SR-capable router MAY allocate more than one Adj-SID to an adjacency.

An SR-capable router MAY allocate the same Adj-SID to different adjacencies.

When the P-Flag is not set, the Adj-SID MAY be persistent. When the P-Flag is set, the Adj-SID MUST be persistent.

Examples of Adj-SID sub-TLV use are described in [RFC8402].

The F-Flag is used in order for the router to advertise the outgoing encapsulation of the adjacency the Adj-SID is attached to.

2.2.2. Adjacency Segment Identifier (LAN-Adj-SID) Sub-TLV

In LAN subnetworks, the Designated Intermediate System (DIS) is elected and originates the Pseudonode LSP (PN LSP) including all neighbors of the DIS.

When Segment Routing is used, each router in the LAN MAY advertise the Adj-SID of each of its neighbors. Since, on LANs, each router only advertises one adjacency to the DIS (and doesn't advertise any other adjacency), each router advertises the set of Adj-SIDs (for each of its neighbors) inside a newly defined sub-TLV that is a part of the TLV advertising the adjacency to the DIS (e.g., TLV-22).

The following new sub-TLV is defined: LAN Adjacency Segment Identifier (LAN-Adj-SID) containing the set of Adj-SIDs the router assigned to each of its LAN neighbors.

The format of the LAN-Adj-SID sub-TLV is as follows:

```

0                               1                               2                               3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type	Length	Flags	Weight
	Neighbor System-ID (ID length octets)		
	SID/Label/Index (variable)		

where:

Type: 32

Length: Variable

Flags: 1-octet field of the following flags:

0	1	2	3	4	5	6	7
F	B	V	L	S	P		

where the F-Flag, B-Flag, V-Flag, L-Flag, S-Flag, and P-Flag are defined in Section 2.2.1.

Other bits: MUST be zero when originated and ignored when received.

Weight: 1 octet. The value represents the weight of the Adj-SID for the purpose of load balancing. The use of the weight is defined in [RFC8402].

Neighbor System-ID: IS-IS System-ID of length "ID Length" as defined in [ISO10589].

SID/Index/Label: As defined in Section 2.1.1.1.

Multiple LAN-Adj-SID sub-TLVs MAY be encoded.

Note that this sub-TLV MUST NOT appear in TLV 141.

In case TLV-22, TLV-23, TLV-222, or TLV-223 (reporting the adjacency to the DIS) can't contain the whole set of LAN-Adj-SID sub-TLVs, multiple advertisements of the adjacency to the DIS MUST be used, and all advertisements MUST have the same metric.

Each router within the level, by receiving the DIS PN LSP as well as the non-PN LSP of each router in the LAN, is capable of reconstructing the LAN topology as well as the set of Adj-SIDs each router uses for each of its neighbors.

2.3. SID/Label Sub-TLV

The SID/Label sub-TLV may be present in the following TLVs/sub-TLVs defined in this document:

SR-Capabilities sub-TLV (Section 3.1)

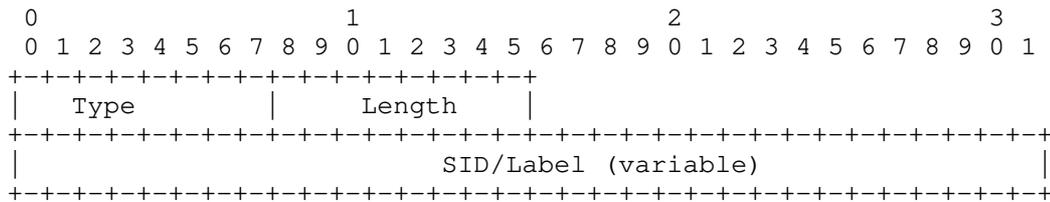
SR Local Block sub-TLV (Section 3.3)

SID/Label Binding TLV (Section 2.4)

Multi-Topology SID/Label Binding TLV (Section 2.5)

Note that the codepoint used in all of the above cases is the SID/Label sub-TLV codepoint specified in the new "sub-TLVs for TLV 149 and 150" registry created by this document.

The SID/Label sub-TLV contains a SID or an MPLS label. The SID/Label sub-TLV has the following format:



where:

Type: 1

Length: 3 or 4

SID/Label: If the length is set to 3, then the 20 rightmost bits represent an MPLS label. If the length is set to 4, then the value is a 32-bit index.

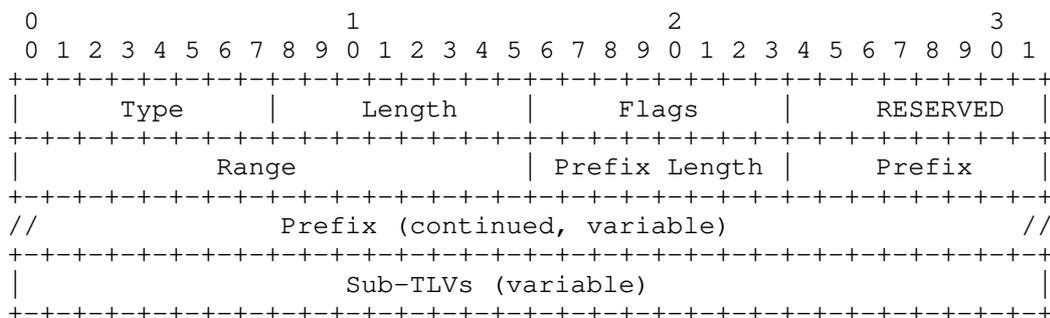
2.4. SID/Label Binding TLV

The SID/Label Binding TLV MAY be originated by any router in an IS-IS domain. There are multiple uses of the SID/Label Binding TLV.

The SID/Label Binding TLV may be used to advertise prefixes to SID/Label mappings. This functionality is called the Segment Routing Mapping Server (SRMS). The behavior of the SRMS is defined in [RFC8661].

The SID/Label Binding TLV may also be used to advertise a Mirror SID indicating the ability of a node to process traffic originally destined to another IGP node. This behavior is defined in [RFC8402].

The SID/Label Binding TLV has the following format:



where:

Type: 149

Length: Variable

Flags: 1 octet

RESERVED: 1 octet (SHOULD be transmitted as 0 and MUST be ignored on receipt)

Range: 2 octets

Prefix Length: 1 octet

Prefix: 0-16 octets

sub-TLVs, where each sub-TLV consists of a sequence of:

- 1 octet of sub-TLV type

- 1 octet of length of the value field of the sub-TLV
- 0-243 octets of value

2.4.1. Flags

Flags: 1-octet field of the following flags:

```

0 1 2 3 4 5 6 7
+---+---+---+---+---+---+
|F|M|S|D|A|   |
+---+---+---+---+---+---+

```

where:

- F-Flag: Address-Family Flag. If unset, then the prefix carries an IPv4 prefix. If set, then the prefix carries an IPv6 prefix.
- M-Flag: Mirror Context Flag. Set if the advertised SID corresponds to a mirrored context. The use of a mirrored context is described in [RFC8402].
- S-Flag: If set, the SID/Label Binding TLV SHOULD be flooded across the entire routing domain. If the S-Flag is not set, the SID/Label Binding TLV MUST NOT be leaked between levels. This bit MUST NOT be altered during the TLV leaking.
- D-Flag: When the SID/Label Binding TLV is leaked from Level-2 to Level-1, the D-Flag MUST be set. Otherwise, this flag MUST be clear. SID/Label Binding TLVs with the D-Flag set MUST NOT be leaked from Level-1 to Level-2. This is to prevent TLV looping across levels.
- A-Flag: Attached Flag. The originator of the SID/Label Binding TLV MAY set the A bit in order to signal that the prefixes and SIDs advertised in the SID/Label Binding TLV are directly connected to their originators. The mechanisms through which the originator of the SID/Label Binding TLV can figure out if a prefix is attached or not are outside the scope of this document (e.g., through explicit configuration). If the Binding TLV is leaked to other areas/levels, the A-Flag MUST be cleared.

An implementation may decide not to honor the S-Flag in order to not leak Binding TLVs between levels (for policy reasons).

Other bits: MUST be zero when originated and ignored when received.

2.4.2. Range

The 'Range' field provides the ability to specify a range of addresses and their associated Prefix-SIDs. This advertisement supports the SRMS functionality. It is essentially a compression scheme to distribute a continuous prefix and their continuous, corresponding SID/Label Block. If a single SID is advertised, then the Range field MUST be set to one. For range advertisements > 1, the Range field MUST be set to the number of addresses that need to be mapped into a Prefix-SID. In either case, the prefix is the first address to which a SID is to be assigned.

2.4.3. Prefix Length, Prefix

The 'Prefix' represents the Forwarding Equivalence Class at the tail end of the advertised path. The 'Prefix' does not need to correspond to a routable prefix of the originating node.

The 'Prefix Length' field contains the length of the prefix in bits. Only the most significant octets of the prefix are encoded (i.e., 1

octet for prefix length 1 up to 8, 2 octets for prefix length 9 to up 16, 3 octets for prefix length 17 up to 24, 4 octets for prefix length 25 up to 32,, and 16 octets for prefix length 113 up to 128).

2.4.4. Mapping Server Prefix-SID

The Prefix-SID sub-TLV is defined in Section 2.1 and contains the SID/Index/Label value associated with the prefix and range. The Prefix-SID sub-TLV MUST be present in the SID/Label Binding TLV when the M-Flag is clear. The Prefix-SID sub-TLV MUST NOT be present when the M-Flag is set.

2.4.4.1. Prefix-SID Flags

The Prefix-SID Flags are defined in Section 2.1. The Mapping Server MAY advertise a mapping with the N-Flag set when the prefix being mapped is known in the link-state topology with a mask length of 32 (IPv4) or 128 (IPv6) and when the prefix represents a node. The mechanisms through which the operator defines that a prefix represents a node are outside the scope of this document (typically it will be through configuration).

The other flags defined in Section 2.1 are not used by the Mapping Server and MUST be ignored at reception.

2.4.4.2. PHP Behavior when Using Mapping Server Advertisements

As the Mapping Server does not specify the originator of a prefix advertisement, it is not possible to determine PHP behavior solely based on the Mapping Server Advertisement. However, if additional information is available, PHP behavior may safely be done. The required information consists of:

- * A prefix reachability advertisement for the prefix has been received, which includes the Prefix Attribute Flags sub-TLV [RFC7794].
- * X-Flag and R-Flag are both set to 0 in the Prefix Attribute Flags sub-TLV.

In the absence of a Prefix Attribute Flags sub-TLV [RFC7794], the A-Flag in the Binding TLV indicates that the originator of a prefix reachability advertisement is directly connected to the prefix; thus, PHP MUST be done by the neighbors of the router originating the prefix reachability advertisement. Note that the A-Flag is only valid in the original area in which the Binding TLV is advertised.

2.4.4.3. Prefix-SID Algorithm

The Algorithm field contains the identifier of the algorithm associated with the SIDs for the prefix(es) in the range. Use of the Algorithm field is described in Section 2.1.

2.4.5. SID/Label Sub-TLV

The SID/Label sub-TLV (Type: 1) contains the SID/Label value as defined in Section 2.3. It MUST be present in the SID/Label Binding TLV when the M-Flag is set in the Flags field of the parent TLV.

2.4.6. Example Encodings

Example 1: If the following IPv4 router addresses (loopback addresses) need to be mapped into the corresponding Prefix-SID indexes, then:

Router-A: 192.0.2.1/32, Prefix-SID: Index 1

Router-B: 192.0.2.2/32, Prefix-SID: Index 2

Router-C: 192.0.2.3/32, Prefix-SID: Index 3

Router-D: 192.0.2.4/32, Prefix-SID: Index 4

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									0	0	0	0	0	0	RESERVED															
Range = 4									32									192																					
0									2									1									Prefix-SID Type												
Sub-TLV Length									Flags									Algorithm																					
																		1																					

Example 2: If the following IPv4 prefixes need to be mapped into the corresponding Prefix-SID indexes, then:

- 10.1.1/24, Prefix-SID: Index 51
- 10.1.2/24, Prefix-SID: Index 52
- 10.1.3/24, Prefix-SID: Index 53
- 10.1.4/24, Prefix-SID: Index 54
- 10.1.5/24, Prefix-SID: Index 55
- 10.1.6/24, Prefix-SID: Index 56
- 10.1.7/24, Prefix-SID: Index 57

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									0	0	0	0	0	0	RESERVED															
Range = 7									24									10																					
1									1									Prefix-SID Type	Sub-TLV Length																				
Flags									Algorithm																														
																		51																					

Example 3: If the following IPv6 prefixes need to be mapped into the corresponding Prefix-SID indexes, then:

- 2001:db8:1/48, Prefix-SID: Index 151
- 2001:db8:2/48, Prefix-SID: Index 152
- 2001:db8:3/48, Prefix-SID: Index 153
- 2001:db8:4/48, Prefix-SID: Index 154

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									1	0	0	0	0	0	RESERVED															
Range = 4									48									0x20																					
0x01									0x0d									0xb8									0x00												
0x01									Prefix-SID Type	Sub-TLV Length									Flags																				
Algorithm																		0																					

```

+-----+
|           151           |
+-----+

```

It is not expected that a network operator will be able to keep fully continuous Prefix/SID/Index mappings. In order to support noncontinuous mapping ranges, an implementation MAY generate several instances of Binding TLVs.

For example, if a router wants to advertise the following ranges:

- Range 16: { 192.0.2.1-15, Index 1-15 }
- Range 6: { 192.0.2.22-27, Index 22-27 }
- Range 41: { 192.0.2.44-84, Index 80-120 }

a router would need to advertise three instances of the Binding TLV.

2.5. Multi-Topology SID/Label Binding TLV

The Multi-Topology SID/Label Binding TLV allows the support of Multi-Topology IS-IS (M-ISIS) as defined in [RFC5120]. The Multi-Topology SID/Label Binding TLV has the same format as the SID/Label Binding TLV defined in Section 2.4 with the difference consisting of a Multi-topology Identifier (MT ID) as defined here below:

```

      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   |           MT ID           |
+-----+-----+-----+-----+
|   Flags   | RESERVED  |           Range           |
+-----+-----+-----+-----+
| Prefix Length | Prefix (variable) | //
+-----+-----+-----+-----+
|           Sub-TLVs (variable)           |
+-----+-----+-----+-----+

```

where:

Type: 150

Length: Variable

MT ID is the Multi-topology Identifier defined as:

```

      0                1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-----+-----+
| RESVD |           MT ID           |
+-----+-----+

```

RESVD: Reserved bits. MUST be reset on transmission and ignored on receive.

MT ID: A 12-bit field containing the non-zero ID of the topology being announced. The TLV MUST be ignored if the ID is zero. This is to ensure the consistent view of the standard unicast topology.

The other fields and sub-TLVs are defined in Section 2.4.

3. Router Capabilities

This section defines sub-TLVs that are inserted into the IS-IS Router Capability that is defined in [RFC7981].

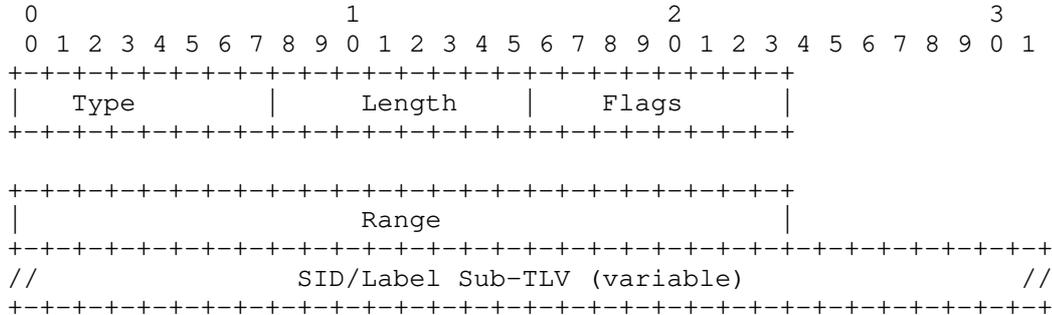
3.1. SR-Capabilities Sub-TLV

Segment Routing requires each router to advertise its SR data plane

capability and the range of MPLS label values it uses for Segment Routing in the case where global SIDs are allocated (i.e., global indexes). Data plane capabilities and label ranges are advertised using the newly defined SR-Capabilities sub-TLV.

The Router Capability TLV specifies flags that control its advertisement. The SR-Capabilities sub-TLV MUST be propagated throughout the level and MUST NOT be advertised across level boundaries. Therefore, Router Capability TLV distribution flags are set accordingly, i.e., the S-Flag in the Router Capability TLV [RFC7981] MUST be unset.

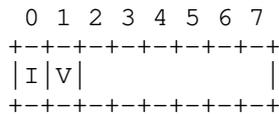
The SR-Capabilities sub-TLV has the following format:



Type: 2

Length: Variable

Flags: 1 octet of flags. The following are defined:



where:

I-Flag: MPLS IPv4 Flag. If set, then the router is capable of processing SR-MPLS-encapsulated IPv4 packets on all interfaces.

V-Flag: MPLS IPv6 Flag. If set, then the router is capable of processing SR-MPLS-encapsulated IPv6 packets on all interfaces.

One or more Segment Routing Global Block (SRGB) Descriptor entries, each of which have the following format:

Range: 3 octets

SID/Label sub-TLV: MPLS label as defined in Section 2.3

The SID/Label sub-TLV contains the first value of the SRGB while the range contains the number of SRGB elements. The range value MUST be higher than 0.

The SR-Capabilities sub-TLV MAY be advertised in an LSP of any number, but a router MUST NOT advertise more than one SR-Capabilities sub-TLV. A router receiving multiple SR-Capabilities sub-TLVs from the same originator SHOULD select the first advertisement in the lowest-numbered LSP.

When multiple SRGB Descriptors are advertised, the entries define an ordered set of ranges on which a SID index is to be applied. For this reason, changing the order in which the descriptors are advertised will have a disruptive effect on forwarding.

When a router adds a new SRGB Descriptor to an existing SR-Capabilities sub-TLV, the new descriptor SHOULD add the newly

configured block at the end of the sub-TLV and SHOULD NOT change the order of previously advertised blocks. Changing the order of the advertised descriptors will create label churn in the FIB and black hole / misdirect some traffic during the IGP convergence. In particular, if a range that is not the last is extended, it's preferable to add a new range rather than extending the previously advertised range.

The originating router MUST ensure the order is unchanged after a graceful restart (using checkpointing, non-volatile storage, or any other mechanism).

The originating router MUST NOT advertise overlapping ranges.

When a router receives multiple overlapping ranges, it MUST conform to the procedures defined in [RFC8660].

Here follows an example of the advertisement of multiple ranges:

The originating router advertises the following ranges:

SR-Cap: range: 100, SID value: 100

SR-Cap: range: 100, SID value: 1000

SR-Cap: range: 100, SID value: 500

The receiving routers concatenate the ranges in the received order and build the SRGB as follows:

```
SRGB = [100, 199]
        [1000, 1099]
        [500, 599]
```

The indexes span multiple ranges:

```
index 0   means label 100
...
index 99  means label 199
index 100 means label 1000
index 199 means label 1099
...
index 200 means label 500
...
```

3.2. SR-Algorithm Sub-TLV

The router may use various algorithms when calculating reachability to other nodes or to prefixes attached to these nodes. Examples of these algorithms are metric-based SPF, various sorts of Constrained SPF, etc. The SR-Algorithm sub-TLV allows the router to advertise the algorithms that the router is currently using. Algorithm values are defined in the "IGP Algorithm Type" registry defined in [RFC8665]. The following values have been defined:

- 0: SPF algorithm based on link metric. This is the well-known shortest path algorithm as computed by the IS-IS Decision Process. Consistent with the deployed practice for link-state protocols, algorithm 0 permits any node to overwrite the SPF path with a different path based on local policy.
- 1: Strict SPF algorithm based on link metric. The algorithm is identical to algorithm 0, but algorithm 1 requires that all nodes along the path will honor the SPF routing decision. Local policy MUST NOT alter the forwarding decision computed by algorithm 1 at the node claiming to support algorithm 1.

The Router Capability TLV specifies flags that control its advertisement. The SR-Algorithm MUST be propagated throughout the level and MUST NOT be advertised across level boundaries. Therefore, Router Capability TLV distribution flags are set accordingly, i.e.,

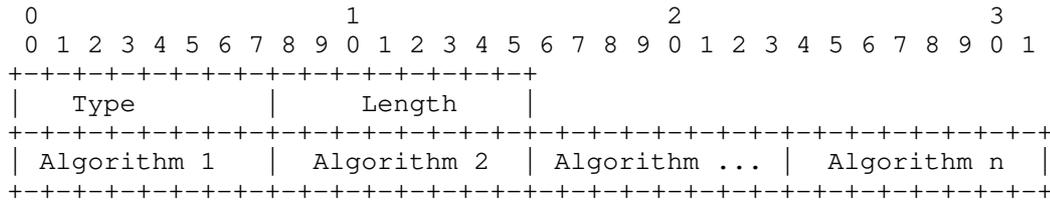
the S-Flag MUST be unset.

The SR-Algorithm sub-TLV is optional. It MUST NOT be advertised more than once at a given level. A router receiving multiple SR-Algorithm sub-TLVs from the same originator SHOULD select the first advertisement in the lowest-numbered LSP.

When the originating router does not advertise the SR-Algorithm sub-TLV, it implies that algorithm 0 is the only algorithm supported by the routers that support the extensions defined in this document.

When the originating router does advertise the SR-Algorithm sub-TLV, then algorithm 0 MUST be present while non-zero algorithms MAY be present.

The SR-Algorithm sub-TLV has the following format:



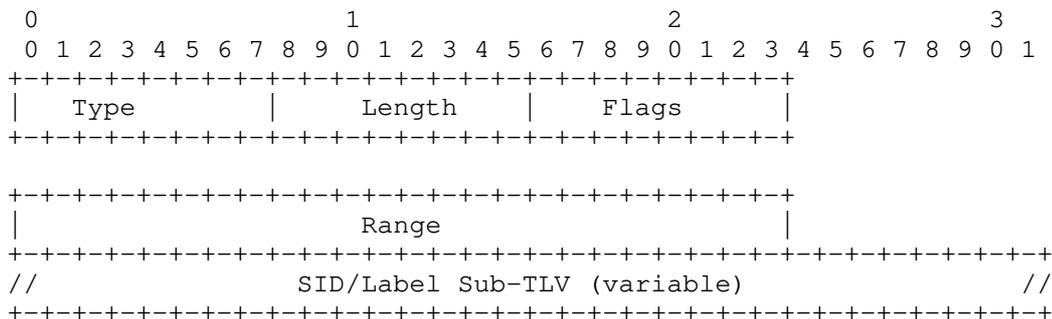
where:

- Type: 19
- Length: Variable
- Algorithm: 1 octet of algorithm

3.3. SR Local Block Sub-TLV

The SR Local Block (SRLB) sub-TLV contains the range of labels the node has reserved for Local SIDs. Local SIDs are used, e.g., for Adj-SIDs, and may also be allocated by components other than the IS-IS protocol. As an example, an application or a controller may instruct the router to allocate a specific Local SID. Therefore, in order for such applications or controllers to know what Local SIDs are available in the router, it is required that the router advertises its SRLB.

The SRLB sub-TLV is used for this purpose and has following format:



- Type: 22
- Length: Variable
- Flags: 1 octet of flags. None are defined at this stage.

One or more SRLB Descriptor entries, each of which have the following format:

- Range: 3 octets
- SID/Label sub-TLV: MPLS label as defined in Section 2.3

31	Adjacency Segment Identifier	y	y	n	y	y	y
32	LAN Adjacency Segment Identifier	y	y	n	y	y	y

Table 1

4.2. Sub-TLVs for Types 135, 235, 236, and 237

This document makes the following registrations in the "Sub-TLVs for TLV 135, 235, 236, and 237" registry.

Type	Description	135	235	236	237
3	Prefix Segment Identifier	y	y	y	y

Table 2

4.3. Sub-TLVs for Type 242

This document makes the following registrations in the "Sub-TLVs for TLV 242" registry.

Type	Description
2	Segment Routing Capability
19	Segment Routing Algorithm
22	Segment Routing Local Block (SRLB)
24	Segment Routing Mapping Server Preference (SRMS Preference)

Table 3

4.4. New TLV Codepoint and Sub-TLV Registry

This document registers the following TLV:

Value	Name	IIH	LSP	SNP	Purge
149	Segment Identifier / Label Binding	n	y	n	n
150	Multi-Topology Segment Identifier / Label Binding	n	y	n	n

Table 4

This document creates the following sub-TLV Registry:

Name: Sub-TLVs for TLVs 149 and 150

Registration Procedure: Expert Review [RFC8126]

Type	Description
0	Reserved
1	SID/Label

2	Unassigned
3	Prefix Segment Identifier
4-255	Unassigned

Table 5

5. Security Considerations

With the use of the extensions defined in this document, IS-IS carries information that will be used to program the MPLS data plane [RFC3031]. In general, the same type of attacks that can be carried out on the IP/IPv6 control plane can be carried out on the MPLS control plane, resulting in traffic being misrouted in the respective data planes. However, the latter may be more difficult to detect and isolate.

Existing security extensions as described in [RFC5304] and [RFC5310] apply to these Segment Routing extensions.

6. References

6.1. Normative References

- [ISO10589] International Organization for Standardization, "Information technology -- Telecommunications and information exchange between systems -- Intermediate system to Intermediate system intra-domain routing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473)", ISO/IEC 10589:2002, Second Edition, November 2002.
- [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>>.
- [RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <<https://www.rfc-editor.org/info/rfc3031>>.
- [RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <<https://www.rfc-editor.org/info/rfc5120>>.
- [RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, DOI 10.17487/RFC5304, October 2008, <<https://www.rfc-editor.org/info/rfc5304>>.
- [RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, DOI 10.17487/RFC5310, February 2009, <<https://www.rfc-editor.org/info/rfc5310>>.
- [RFC7794] Ginsberg, L., Ed., Decraene, B., Previdi, S., Xu, X., and U. Chunduri, "IS-IS Prefix Attributes for Extended IPv4 and IPv6 Reachability", RFC 7794, DOI 10.17487/RFC7794, March 2016, <<https://www.rfc-editor.org/info/rfc7794>>.
- [RFC7981] Ginsberg, L., Previdi, S., and M. Chen, "IS-IS Extensions for Advertising Router Information", RFC 7981, DOI 10.17487/RFC7981, October 2016, <<https://www.rfc-editor.org/info/rfc7981>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC

2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.
- [RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <<https://www.rfc-editor.org/info/rfc8660>>.
- [RFC8661] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., and S. Litkowski, "Segment Routing MPLS Interworking with LDP", RFC 8661, DOI 10.17487/RFC8661, December 2019, <<https://www.rfc-editor.org/info/rfc8661>>.
- [RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019, <<https://www.rfc-editor.org/info/rfc8665>>.

6.2. Informative References

- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <<https://www.rfc-editor.org/info/rfc5308>>.
- [RFC5311] McPherson, D., Ed., Ginsberg, L., Previdi, S., and M. Shand, "Simplified Extension of Link State PDU (LSP) Space for IS-IS", RFC 5311, DOI 10.17487/RFC5311, February 2009, <<https://www.rfc-editor.org/info/rfc5311>>.
- [RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, DOI 10.17487/RFC5316, December 2008, <<https://www.rfc-editor.org/info/rfc5316>>.
- [RFC7855] Previdi, S., Ed., Filsfils, C., Ed., Decraene, B., Litkowski, S., Horneffer, M., and R. Shakir, "Source Packet Routing in Networking (SPRING) Problem Statement and Requirements", RFC 7855, DOI 10.17487/RFC7855, May 2016, <<https://www.rfc-editor.org/info/rfc7855>>.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.

Acknowledgements

We would like to thank Dave Ward, Dan Frost, Stewart Bryant, Pierre Francois, and Jesper Skriver for their contribution to the content of this document.

Contributors

The following people gave a substantial contribution to the content of this document and should be considered as coauthors:

Stephane Litkowski
Orange
France
Email: stephane.litkowski@orange.com

Jeff Tantsura
Apstra, Inc.
Email: jefftant@gmail.com

Peter Psenak
Cisco Systems Inc.
United States of America
Email: ppsenak@cisco.com

Martin Horneffer
Deutsche Telekom
Germany
Email: Martin.Horneffer@telekom.de

Wim Henderickx
Nokia
Belgium
Email: wim.henderickx@nokia.com

Edward Crabbe
Oracle
United States of America
Email: edward.crabbe@oracle.com

Rob Shakir
Google
United Kingdom
Email: robjs@google.com

Igor Milojevic
Individual
Serbia
Email: milojevicigor@gmail.com

Saku Ytti
TDC
Finland
Email: saku@ytti.fi

Authors' Addresses

Stefano Previdi (editor)
Huawei Technologies
Italy
Email: stefano@previdi.net

Les Ginsberg (editor)
Cisco Systems, Inc.
United States of America
Email: ginsberg@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Brussels
Belgium
Email: cfilsfil@cisco.com

Ahmed Bashandy
Arrcus
Email: abashandy.ietf@gmail.com

Hannes Gredler

RtBrick Inc.

Email: hannes@rtbrick.com

Bruno Decraene

Orange

France

Email: bruno.decraene@orange.com