Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: February 1, 2016

C. Filsfils, Ed.
S. Previdi, Ed.
Cisco Systems, Inc.
B. Decraene
S. Litkowski
Orange
R. Shakir
Individual
July 31, 2015

# Segment Routing Architecture
## draft-ietf-spring-segment-routing-04

Abstract

Segment Routing (SR) leverages the source routing paradigm. A node steers a packet through an ordered list of instructions, called segments. A segment can represent any instruction, topological or service-based. A segment can have a local semantic to an SR node or global within an SR domain. SR allows to enforce a flow through any topological path and service chain while maintaining per-flow state only at the ingress node to the SR domain.

Segment Routing can be directly applied to the MPLS architecture with no change on the forwarding plane. A segment is encoded as an MPLS label. An ordered list of segments is encoded as a stack of labels. The segment to process is on the top of the stack. Upon completion of a segment, the related label is popped from the stack.

Segment Routing can be applied to the IPv6 architecture, with a new type of routing extension header. A segment is encoded as an IPv6 address. An ordered list of segments is encoded as an ordered list of IPv6 addresses in the routing extension header. The segment to process is indicated by a pointer in the routing extension header. Upon completion of a segment, the pointer is incremented.

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].

This Internet-Draft will expire on February 1, 2016.

Copyright Notice

Table of Contents

## 1. Introduction

With Segment Routing (SR), a node steers a packet through an ordered list of instructions, called segments. A segment can represent any instruction, topological or service-based. A segment can have a local semantic to an SR node or global within an SR domain. SR allows to enforce a flow through any path and service chain while maintaining per-flow state only at the ingress node of the SR domain.

Segment Routing can be directly applied to the MPLS architecture (RFC 3031) with no change on the forwarding plane. A segment is encoded as an MPLS label. An ordered list of segments is encoded as a stack of labels. The active segment is on the top of the stack. A completed segment is popped off the stack. The addition of a segment is performed with a push.

In the Segment Routing MPLS instantiation, a segment could be of several types:

- an IGP segment,
- a BGP Peering segments,
- an LDP LSP segment,
- an RSVP-TE LSP segment,
- a BGP LSP segment.

The first two (IGP and BGP Peering segments) types of segments defined in this document. The use of the last three types of segments is illustrated in [I-D.ietf-spring-segment-routing-mpls].

Segment Routing can be applied to the IPv6 architecture (RFC2460), with a new type of routing extension header. A segment is encoded as an IPv6 address. An ordered list of segments is encoded as an ordered list of IPv6 addresses in the routing extension header. The active segment is indicated by a pointer in the routing extension header. Upon completion of a segment, the pointer is incremented. A segment can be inserted in the list and the pointer is updated accordingly.

Numerous use-cases illustrate the benefits of source routing either for FRR, OAM or Traffic Engineering reasons.

This document defines a set of instructions (called segments) that are required to fulfill the described use-cases. These segments can either be used in isolation (one single segment defines the source

route of the packet) or in combination (these segments are part of an ordered list of segments that define the source route of the packet).

### 1.1. Companion Documents

This document defines the SR architecture, its routing model, the IGP-based segments, the BGP-based segments and the service segments.

Use cases are described in [I-D.ietf-spring-problem-statement], [I-D.filsfils-spring-segment-routing-central-epe], [I-D.filsfils-spring-segment-routing-msdc], [I-D.ietf-spring-ipv6-use-cases], [I-D.ietf-spring-resiliency-use-cases], [I-D.geib-spring-oam-usecase] and [I-D.kumar-spring-sr-oam-requirement].

Segment Routing for MPLS dataplane is documented in [I-D.ietf-spring-segment-routing-mpls].

Segment Routing for IPv6 dataplane is documented in [I-D.previdi-6man-segment-routing-header].

IGP protocol extensions for Segment Routing are described in [I-D.ietf-isis-segment-routing-extensions], [I-D.ietf-ospf-segment-routing-extensions] and [I-D.ietf-ospf-ospfv3-segment-routing-extensions] referred in this document as "IGP SR extensions documents".

The FRR solution for SR is documented in [I-D.francois-spring-segment-routing-ti-lfa].

The PCEP protocol extensions for Segment Routing are defined in [I-D.ietf-pce-segment-routing].

The interaction between SR/MPLS with other MPLS Signaling planes is documented in [I-D.filsfils-spring-segment-routing-ldp-interop].

## 2. Terminology

Segment: an instruction a node executes on the incoming packet (e.g.: forward packet according to shortest path to destination, or, forward packet through a specific interface, or, deliver the packet to a given application/service instance).

SID: a Segment Identifier

Segment List: ordered list of SID's encoding the topological and service source route of the packet. It is a stack of labels in the

MPLS architecture. It is an ordered list of IPv6 addresses in the IPv6 architecture.

Active segment: the segment that MUST be used by the receiving router to process the packet. It is identified by a pointer in the IPv6 architecture. It is the top label in the MPLS architecture.

PUSH: the insertion of a segment at the head of the Segment list.

NEXT: the active segment is completed, the next segment becomes active.

CONTINUE: the active segment is not completed and hence remains active. The CONTINUE instruction is implemented as the SWAP instruction in the MPLS dataplane.

SR Global Block (SRGB): local property of an SR node. In the MPLS architecture, SRGB is the set of local labels reserved for global segments. In the IPv6 architecture, it is the set of locally relevant IPv6 addresses. Using the same SRGB on all nodes within the SR domain ease operations and troubleshooting and is expected to be a deployment guideline.

Global Segment: the related instruction is supported by all the SR-capable nodes in the domain. In the MPLS architecture, a Global Segment has a globally-unique index. The related local label at a given node N is found by adding the globally-unique index to the SRGB of node N. In the IPv6 architecture, a global segment is a globally-unique IPv6 address.

Local Segment: the related instruction is supported only by the node originating it. In the MPLS architecture, this is a local label outside the SRGB. In the IPv6 architecture, this is a link-local address.

IGP Segment: the generic name for a segment attached to a piece of information advertised by a link-state IGP, e.g. an IGP prefix or an IGP adjacency.

IGP-prefix Segment, Prefix-SID: an IGP-Prefix Segment is an IGP segment attached to an IGP prefix. An IGP-Prefix Segment is always global within the SR/IGP domain and identifies an instruction to forward the packet over the ECMP-aware shortest-path computed by the IGP to the related prefix. The Prefix-SID is the SID of the IGP-Prefix Segment.

IGP-Anycast: an IGP-Anycast Segment is an IGP-prefix segment which does not identify a specific router, but a set of routers. The terms

"Anycast Segment" or "Anycast-SID" are often used as an abbreviation.

IGP-Adjacency: an IGP-Adjacency Segment is an IGP segment attached to an unidirectional adjacency or a set of unidirectional adjacencies. By default, an IGP-Adjacency Segment is local (unless explicitly advertised otherwise) to the node that advertises it.

IGP-Node: an IGP-Node Segment is an IGP-Prefix Segment which identifies a specific router (e.g. a loopback). The terms "Node Segment" or Node-SID" are often used as an abbreviation.

SR Tunnel: a list of segments to be pushed on the packets directed on the tunnel. The list of segments can be specified explicitly or implicitly via a set of abstract constraints (latency, affinity, SRLG, ...). In the latter case, a constraint-based path computation is used to determine the list of segments associated with the tunnel. The computation can be local or delegated to a PCE server. An SR tunnel can be configured by the operator, provisioned via netconf or provisioned via PCEP. An SR tunnel can be used for traffic-engineering, OAM or FRR reasons.

Segment List Depth: the number of segments of an SR tunnel. The entity instantiating an SR Tunnel at a node N should be able to discover the depth insertion capability of the node N. The PCEP discovery capability is described in [I-D.ietf-pce-segment-routing].

## 3. Link-State IGP Segments

Within a link-state IGP domain, an SR-capable IGP node advertises segments for its attached prefixes and adjacencies. These segments are called IGP segments or IGP SIDs. They play a key role in Segment Routing and use-cases as they enable the expression of any topological path throughout the IGP domain. Such a topological path is either expressed as a single IGP segment or a list of multiple IGP segments.

### 3.1. IGP Segment, IGP SID

The terms "IGP Segment" and "IGP SID" are the generic names for a segment attached to a piece of information advertised by a link-state IGP, e.g. an IGP prefix or an IGP adjacency.

### 3.2. IGP-Prefix Segment, Prefix-SID

An IGP-Prefix Segment is an IGP segment attached to an IGP prefix. An IGP-Prefix Segment is always global within the SR/IGP domain and identifies the ECMP-aware shortest-path computed by the IGP to the

related prefix. The Prefix-SID is the SID of the IGP-Prefix Segment.

A packet injected anywhere within the SR/IGP domain with an active Prefix-SID will be forwarded along the shortest-path to that prefix.

The IGP signaling extension for IGP-Prefix segment includes a set of flags. The encoding details of the Prefix-SID and its flags are described in IGP SR extensions documents.

The IGP signaling extension for IGP-Prefix segment includes the P-Flag. A Node N advertising a Prefix-SID SID-R for its attached prefix R resets the P-Flag to allow its connected neighbors to perform the NEXT operation while processing SID-R. This behavior is equivalent to Penultimate Hop Popping in MPLS. When set, the neighbors of N must perform the CONTINUE operation while processing SID-R.

While SR allows to attach a local segment to an IGP prefix (using the L-Flag), we specifically assume that when the terms "IGP-Prefix Segment" and "Prefix-SID" are used, the segment is global (the SID is allocated from the SRGB). This is consistent with all the described use-cases that require global segments attached to IGP prefixes.

A single Prefix-SID is allocated to an IGP Prefix in a topology.

In the context of multiple topologies, multiple Prefix-SID's MAY be allocated to the same IGP Prefix (e.g.: using the "algorithm" field in the IGP advertisement as described in IGP SR extensions documents). However, each prefix-SID MUST be associated with only one topology. In other words: a prefix, within a topology, MUST have only a single Prefix-SID.

A Prefix-SID is allocated from the SRGB according to a process similar to IP address allocation. Typically the Prefix-SID is allocated by policy by the operator (or NMS) and the SID very rarely changes.

The allocation process MUST NOT allocate the same Prefix-SID to different IP prefixes.

If a node learns a Prefix-SID having a value that falls outside the locally configured SRGB range, then the node MUST NOT use the Prefix-SID and SHOULD issue an error log warning for misconfiguration.

The required IGP protocol extensions are defined in IGP SR extensions documents.

A node N attaching a Prefix-SID SID-R to its attached prefix R MUST

maintain the following FIB entry:

```
   Incoming Active Segment: SID-R
   Ingress Operation: NEXT
   Egress interface: NULL
```

A remote node M MUST maintain the following FIB entry for any learned Prefix-SID SID-R attached to IP prefix R:

```
   Incoming Active Segment: SID-R
   Ingress Operation:
      If the next-hop of R is the originator of R
      and instructed to remove the active segment: NEXT
      Else: CONTINUE
   Egress interface: the interface towards the next-hop along
                     the shortest-path to prefix R.
```

### 3.3. IGP-Node Segment, Node-SID

An IGP-Node Segment is a an IGP-Prefix Segment which identifies a specific router (e.g. a loopback). The N flag is set. The terms "Node Segment" or "Node-SID" are often used as an abbreviation.

A "Node Segment" or "Node-SID" is fundamental to SR. From anywhere in the network, it enforces the ECMP-aware shortest-path forwarding of the packet towards the related node.

An IGP-Node-SID MUST NOT be associated with a prefix that is owned or advertised by more than one router within the same routing domain.

### 3.4. IGP-Anycast Segment, Anycast SID

An IGP-Anycast Segment is an IGP-prefix segment which does not identify a specific router, but a set of routers. The terms "Anycast Segment" or "Anycast-SID" are often used as an abbreviation.

An "Anycast Segment" or "Anycast SID" enforces the ECMP-aware shortest-path forwarding towards the closest node of the anycast set. This is useful to express macro-engineering policies or protection mechanisms.

An IGP-Anycast Segment MUST NOT reference a particular node.

Within an anycast group, all routers MUST advertise the same prefix with the same SID value.

```
                               +--------------+
                               |   Group A    |
                               | 192.0.1.1/32 |
                               |    SID:100   |
                               |              |
                        +-----------A1---A3----------+
                        |      |    | \ / |   |      |
             SID:10     |      |    |  /  |   |      |     SID:30
           1.1.1.1/32   |      |    | / \ |   |      |   1.1.1.3/32
               PE1------R1----------A2---A4---------R3------PE3
                 \     /|      |              |      |\     /
                  \   / |      +--------------+      | \   /
                   \ /  |                            |  \ /
                    /   |                            |   /
                   / \  |                            |  / \
                  /   \ |      +--------------+      | /   \
                 /     \|      |              |      |/     \
               PE2------R2----------B1---B3----+----R4------PE4
           1.1.1.2/32   |      |    | \ / |   |      |   1.1.1.4/32
             SID:20     |      |    |  /  |   |      |     SID:40
                        |      |    | / \ |   |      |
                        +-----+-----B2---B4----+-----+
                               |              |
                               |   Group B    |
                               | 192.0.2.1/32 |
                               |    SID:200   |
                               +--------------+
```

Transit device groups

The figure above describes a network example with two groups of transit devices. Group A consists of devices {A1, A2, A3 and A4}. They are all provisioned with the anycast address 192.0.1.1/32 and the anycast SID 100.

Similarly, group B consists of devices {B1, B2, B3 and B4} and are all provisioned with the anycast address 192.0.1.2/32, anycast SID 200. In the above network topology, each PE device is connected to two routers in each of the groups A and B.

PE1 can choose a particular transit device group when sending traffic to PE3 or PE4. This will be done by pushing the anycast SID of the group in the stack.

Processing the anycast, and subsequent segments, requires special care.

Obviously, the value of the SID following the anycast SID MUST be

understood by all nodes advertising the same anycast segment.

```
                         +-------------------------+
                         |       Group A           |
                         |     192.0.1.1/32        |
                         |        SID:100          |
                         |-------------------------|
                         |                         |
                         |   SRGB:         SRGB:   |
      SID:10             |(1000-2000)   (3000-4000)|             SID:30
        PE1---+       +-------A1-------------A3-------+       +---PE3
               \     /   |    | \           / |    |   \     /
                \   /    |    |  +-----+   /  |    |    \   /
         SRGB:   \ /     |    |         \ /   |    |     \ /   SRGB:
      (7000-8000) R1     |    |          \    |    |      R3 (6000-7000)
                 / \     |    |         / \   |    |     / \
                /   \    |    |  +-----+   \  |    |    /   \
               /     \   |    | /           \ |    |   /     \
        PE2---+       +-------A2-------------A4-------+       +---PE4
      SID:20             |   SRGB:         SRGB:   |             SID:40
                         |(2000-3000)   (4000-5000)|
                         |                         |
                         +-------------------------+
```

Transit paths via anycast group A

Considering a MPLS deployment, in the above topology, if device PE1 (or PE2) requires to send a packet to the device PE3 (or PE4) it needs to encapsulate the packet in a MPLS payload with the following stack of labels.

- o Label allocated by R1 for anycast SID 100 (outer label).

- o Label allocated by the nearest router in group A for SID 30 (for destination PE3).

While the first label is easy to compute, in this case since there are more than one topologically nearest devices (A1 and A2), unless A1 and A2 allocated the same label value to the same prefix, determining the second label is impossible. Devices A1 and A2 may be devices from different hardware vendors. If both don't allocate the same label value for SID 30, it is impossible to use the anycast group "A" as a transit anycast group towards PE3. Hence, PE1 (or PE2) cannot compute an appropriate label stack to steer the packet exclusively through the group A devices. Same holds true for devices PE3 and PE4 when trying to send a packet to PE1 or PE2.

To ease the use of anycast segment in a short term, it is recommended to configure the same SRGB on all nodes of a particular anycast

group. Using this method, as mentioned above, computation of the label following the anycast segment is straightforward.

Using anycast segment without configuring the same SRGB on nodes belonging to the same device group may lead to misrouting (in a MPLS VPN deployment, some traffic may leak between VPNs).

### 3.5. IGP-Adjacency Segment, Adj-SID

An IGP-Adjacency Segment is an IGP segment attached to a unidirectional adjacency or a set of unidirectional adjacencies. By default, an IGP-Adjacency Segment is local to the node which advertises it. However, an Adjacency Segment can be global if advertised by the IGP as such. The SID of the IGP-Adjacency Segment is called the Adj-SID.

The adjacency is formed by the local node (i.e., the node advertising the adjacency in the IGP) and the remote node (i.e., the other end of the adjacency). The local node MUST be an IGP node. The remote node MAY be an adjacent IGP neighbor) or a non-adjacent neighbor (e.g.: a Forwarding Adjacency, [RFC4206]).

A packet injected anywhere within the SR domain with a segment list {SN, SNL}, where SN is the Node-SID of node N and SNL is an Adj-SID attached by node N to its adjacency over link L, will be forwarded along the shortest-path to N and then be switched by N, without any IP shortest-path consideration, towards link L. If the Adj-SID identifies a set of adjacencies, then the node N load- balances the traffic among the various members of the set.

Similarly, when using a global Adj-SID, a packet injected anywhere within the SR domain with a segment list {SNL}, where SNL is a global Adj-SID attached by node N to its adjacency over link L, will be forwarded along the shortest-path to N and then be switched by N, without any IP shortest-path consideration, towards link L. If the Adj-SID identifies a set of adjacencies, then the node N load-balances the traffic among the various members of the set. The use of global Adj-SID allows to reduce the size of the segment list when expressing a path at the cost of additional state (i.e.: the global Adj-SID will be inserted by all routers within the area in their forwarding table).

An "IGP Adjacency Segment" or "Adj-SID" enforces the switching of the packet from a node towards a defined interface or set of interfaces. This is key to theoretically prove that any path can be expressed as a list of segments.

The encodings of the Adj-SID include the B-flag. When set, the Adj-

SID benefits from a local protection.

The encodings of the Adj-SID include the L-flag. When set, the Adj-SID has local significance. By default the L-flag is set.

A node SHOULD allocate one Adj-SIDs for each of its adjacencies.

A node MAY allocate multiple Adj-SIDs to the same adjacency. An example is where the adjacency is established over a bundle interface. Each bundle member MAY have its own Adj-SID.

A node MAY allocate the same Adj-SID to multiple adjacencies.

Adjacency suppression MUST NOT be performed by the IGP.

A node MUST install a FIB entry for any Adj-SID of value V attached to data-link L:

```
   Incoming Active Segment: V
   Operation: NEXT
   Egress Interface: L
```

The Adj-SID implies, from the router advertising it, the forwarding of the packet through the adjacency identified by the Adj-SID, regardless its IGP/SPF cost. In other words, the use of Adjacency Segments overrides the routing decision made by SPF algorithm.

### 3.5.1. Parallel Adjacencies

Adj-SIDs can be used in order to represent a set of parallel interfaces between two adjacent routers.

A node MUST install a FIB entry for any locally originated Adjacency Segment (Adj-SID) of value W attached to a set of link B with:

```
   Incoming Active Segment: W
   Ingress Operation: NEXT
   Egress interface: loadbalance between any data-link within set B
```

When parallel adjacencies are used and associated to the same Adj-SID, and in order to optimize the load balancing function, a "weight" factor can be associated to the Adj-SID advertised with each adjacency. The weight tells the ingress (or a SDN/orchestration system) about the loadbalancing factor over the parallel adjacencies. As shown in Figure 1, A and B are connected through two parallel adjacencies

```
                   link-1
                 +--------+
                 |        |
             S---A        B---C
                 |        |
                 +--------+
                   link-2
```

Figure 1: Parallel Links and Adj-SIDs

Node A advertises following Adj-SIDs and weights:

- o Link-1: Adj-SID 1000, weight: 1

- o Link-2: Adj-SID 1000, weight: 2

Node S receives the advertisements of the parallel adjacencies and understands that by using Adj-SID 1000 node A will loadbalance the traffic across the parallel links (link-1 and link-2) according to a 1:2 ratio.

The weight value is advertised with the Adj-SID as defined in IGP SR extensions documents.

### 3.5.2. LAN Adjacency Segments

In LAN subnetworks, link-state protocols define the concept of Designated Router (DR, in OSPF) or Designated Intermediate System (DIS, in IS-IS) that conduct flooding in broadcast subnetworks and that describe the LAN topology in a special routing update (OSPF Type2 LSA or IS-IS Pseudonode LSP).

The difficulty with LANs is that each router only advertises its connectivity to the DR/DIS and not to each other individual nodes in the LAN. Therefore, additional protocol mechanisms (IS-IS and OSPF) are necessary in order for each router in the LAN to advertise an Adj-SID associated to each neighbor in the LAN. These extensions are defined in IGP SR extensions documents.

## 3.6. Binding Segment

### 3.6.1. Mapping Server

A Remote-Binding SID S advertised by the mapping server M for remote prefix R attached to non-SR-capable node N signals the same information as if N had advertised S as a Prefix-SID. Further details are described in the SR/LDP interworking procedures ([I-D.filsfils-spring-segment-routing-ldp-interop].

The segment allocation and SRGB Maintenance rules are the same as those defined for Prefix-SID.

### 3.6.2. Tunnel Headend

The segment allocation and SRGB Maintenance rules are the same as those defined for Adj-SID. A tunnel attached to a head-end H acts as an adjacency attached to H.

Note: an alternative would consist in representing tunnels as forwarding-adjacencies ( [RFC4206]). In such case, the tunnel is presented to the routing area as a routing adjacency and will be considered as such by all area routers. The Remote-Binding SID is preferred as it allows to advertise the presence of a tunnel without influencing the LSDB and the SPF computation.

## 3.7. Inter-Area Considerations

In the following example diagram we assume an IGP deployed using areas and where SR has been deployed.

```
               !          !
               !          !
        B------C-----F----G-----K
       /       |          |     |
 S---A/        |          |     |
      \        |          |     |
       \D------I----------J-----L----Z (192.0.2.1/32, Node-SID: 150)
               !          !
       Area-1  ! Backbone ! Area 2
               !   area   !
```

Figure 2: Inter-Area Topology Example

In area 2, node Z allocates Node-SID 150 to his local prefix 192.0.2.1/32. ABRs G and J will propagate the prefix into the backbone area by creating a new instance of the prefix according to normal inter-area/level IGP propagation rules.

Nodes C and I will apply the same behavior when leaking prefixes from the backbone area down to area 1. Therefore, node S will see prefix 192.0.2.1/32 with Prefix-SID 150 and advertised by nodes C and I.

It therefore results that a Prefix-SID remains attached to its related IGP Prefix through the inter-area process.

When node S sends traffic to 192.0.2.1/32, it pushes Node-SID(150) as active segment and forward it to A.

When packet arrives at ABR I (or C), the ABR forwards the packet according to the active segment (Node-SID(150)). Forwarding continues across area borders, using the same Node-SID(150), until the packet reaches its destination.

When an ABR propagates a prefix from one area to another it MUST set the R-Flag.

## 4. BGP Peering Segments

In the context of BGP Egress Peer Engineering (EPE), as described in [I-D.filsfils-spring-segment-routing-central-epe], an EPE enabled Egress PE node MAY advertise segments corresponding to its attached peers. These segments are called BGP peering segments or BGP Peering SIDs. They enable the expression of source-routed inter-domain paths.

An ingress border router of an AS may compose a list of segments to steer a flow along a selected path within the AS, towards a selected egress border router C of the AS and through a specific peer. At minimum, a BGP Peering Engineering policy applied at an ingress PE involves two segments: the Node SID of the chosen egress PE and then the BGP Peering Segment for the chosen egress PE peer or peering interface.

Hereafter, we will define three types of BGP peering segments/SID's: PeerNodeSID, PeerAdjSID and PeerSetSID.

- o PeerNode SID. A BGP PeerNode segment/SID is a local segment. At the BGP node advertising it, its semantics is:

  - * SR header operation: NEXT.

  - * Next-Hop: the connected peering node to which the segment is related.

- o PeerAdj SID: A BGP PeerAdj segment/SID is a local segment. At the BGP node advertising it, its semantics is:

  - * SR header operation: NEXT.

  - * Next-Hop: the peer connected through the interface to which the segment is related.

- o PeerSet SID. A BGP PeerSet segment/SID is a local segment. At the BGP node advertising it, its semantics is:

* SR header operation: NEXT.

* Next-Hop: loadbalance across any connected interface to any peer in the related group.

A peer set could be all the connected peers from the same AS or a subset of these. A group could also span across AS. The group definition is a policy set by the operator.

The BGP extensions necessary in order to signal these BGP peering segments will be defined in a separate document.

## 5. Multicast

Segment Routing is defined for unicast. The application of the source-route concept to Multicast is not in the scope of this document.

## 6. IANA Considerations

This document does not require any action from IANA.

## 7. Security Considerations

This document doesn't introduce new security considerations when applied to the MPLS dataplane.

There are a number of security concerns with source routing at the IPv6 dataplane [RFC5095]. The new IPv6-based segment routing header defined in [I-D.previdi-6man-segment-routing-header] and its associated security measures address these concerns. The IPv6 Segment Routing Header is defined in a way that blind attacks are never possible, i.e., attackers will be unable to send source routed packets that get successfully processed, without being part of the negations for setting up the source routes or being able to eavesdrop legitimate source routed packets. In some networks this base level security may be complemented with other mechanisms, such as packet filtering, cryptographic security, etc.

## 8. Contributors

The following people have substantially contributed to the definition of the Segment Routing architecture and to the editing of this document:

Ahmed Bashandy
Cisco Systems, Inc.
Email: bashandy@cisco.com
Martin Horneffer
Deutsche Telekom
Email: Martin.Horneffer@telekom.de
Wim Henderickx
Alcatel-Lucent
Email: wim.henderickx@alcatel-lucent.com
Jeff Tantsura
Ericsson
Email: Jeff.Tantsura@ericsson.com
Edward Crabbe
Individual
Email: edward.crabbe@gmail.com
Igor Milojevic
Email: milojevicigor@gmail.com
Saku Ytti
TDC
Email: saku@ytti.fi

## 9. Acknowledgements

We would like to thank Dave Ward, Dan Frost, Stewart Bryant, Pierre Francois, Thomas Telkamp, Les Ginsberg, Ruediger Geib, Hannes Gredler and Pushpasis Sarkar for their comments and review of this document.

## 10. References

### 10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, DOI 10.17487/RFC4206, October 2005, <http://www.rfc-editor.org/info/rfc4206>.

### 10.2. Informative References

[I-D.filsfils-spring-segment-routing-central-epe]
Filsfils, C., Previdi, S., Patel, K., Aries, E.,

shaw@fb.com, s., Ginsburg, D., and D. Afanasiev, "Segment Routing Centralized Egress Peer Engineering", draft-filsfils-spring-segment-routing-central-epe-04 (work in progress), July 2015.

[I-D.filsfils-spring-segment-routing-ldp-interop]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing interoperability with LDP", draft-filsfils-spring-segment-routing-ldp-interop-03 (work in progress), March 2015.

[I-D.filsfils-spring-segment-routing-msdc]
Filsfils, C., Previdi, S., Mitchell, J., Aries, E., Lapukhov, P., Gaya, G., Afanasiev, D., Laberge, T., Nkposong, E., Nanduri, M., Uttaro, J., and S. Ray, "BGP-Prefix Segment in large-scale data centers", draft-filsfils-spring-segment-routing-msdc-03 (work in progress), July 2015.

[I-D.francois-spring-segment-routing-ti-lfa]
Francois, P., Filsfils, C., Bashandy, A., and B. Decraene, "Topology Independent Fast Reroute using Segment Routing", draft-francois-spring-segment-routing-ti-lfa-01 (work in progress), October 2014.

[I-D.geib-spring-oam-usecase]
Geib, R., Filsfils, C., Pignataro, C., and N. Kumar, "Use case for a scalable and topology aware MPLS data plane monitoring system", draft-geib-spring-oam-usecase-06 (work in progress), July 2015.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and J. Tantsura, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-05 (work in progress), June 2015.

[I-D.ietf-ospf-ospfv3-segment-routing-extensions]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPFv3 Extensions for Segment Routing", draft-ietf-ospf-ospfv3-segment-routing-extensions-03 (work in progress), June 2015.

[I-D.ietf-ospf-segment-routing-extensions]

Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", draft-ietf-ospf-segment-routing-extensions-05 (work in progress), June 2015.

[I-D.ietf-pce-segment-routing]
Sivabalan, S., Medved, J., Filsfils, C., Crabbe, E., Lopez, V., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", draft-ietf-pce-segment-routing-05 (work in progress), May 2015.

[I-D.ietf-spring-ipv6-use-cases]
Brzozowski, J., Leddy, J., Leung, I., Previdi, S., Townsley, W., Martin, C., Filsfils, C., and R. Maglione, "IPv6 SPRING Use Cases", draft-ietf-spring-ipv6-use-cases-04 (work in progress), March 2015.

[I-D.ietf-spring-problem-statement]
Previdi, S., Filsfils, C., Decraene, B., Litkowski, S., Horneffer, M., and R. Shakir, "SPRING Problem Statement and Requirements", draft-ietf-spring-problem-statement-04 (work in progress), April 2015.

[I-D.ietf-spring-resiliency-use-cases]
Francois, P., Filsfils, C., Decraene, B., and R. Shakir, "Use-cases for Resiliency in SPRING", draft-ietf-spring-resiliency-use-cases-01 (work in progress), March 2015.

[I-D.ietf-spring-segment-routing-mpls]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Shakir, R., Tantsura, J., and E. Crabbe, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls-01 (work in progress), May 2015.

[I-D.kumar-spring-sr-oam-requirement]
Kumar, N., Pignataro, C., Akiya, N., Geib, R., Mirsky, G., and S. Litkowski, "OAM Requirements for Segment Routing Network", draft-kumar-spring-sr-oam-requirement-03 (work in progress), March 2015.

[I-D.previdi-6man-segment-routing-header]
Previdi, S., Filsfils, C., Field, B., Leung, I., Aries, E., Vyncke, E., and D. Lebrun, "IPv6 Segment Routing

Header (SRH)",
draft-previdi-6man-segment-routing-header-07 (work in
progress), July 2015.

[RFC5095] Abley, J., Savola, P., and G. Neville-Neil, "Deprecation
of Type 0 Routing Headers in IPv6", RFC 5095,
DOI 10.17487/RFC5095, December 2007,
<http://www.rfc-editor.org/info/rfc5095>.

## Authors' Addresses

Clarence Filsfils (editor)
Cisco Systems, Inc.
Brussels,
BE

Email: cfilsfil@cisco.com

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

Email: sprevidi@cisco.com

Bruno Decraene
Orange
FR

Email: bruno.decraene@orange.com

Stephane Litkowski
Orange
FR

Email: stephane.litkowski@orange.com

Rob Shakir
Individual

Email: rjs@rob.sh