Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 23, 2014

C. Filsfils, Ed.
S. Previdi, Ed.
A. Bashandy
Cisco Systems, Inc.
B. Decraene
S. Litkowski
Orange
M. Horneffer
Deutsche Telekom
I. Milojevic
Telekom Srbija
R. Shakir
British Telecom
S. Ytti
TDC Oy
W. Henderickx
Alcatel-Lucent
J. Tantsura
Ericsson
E. Crabbe
Google, Inc.
October 20, 2013

Segment Routing with MPLS data plane
draft-filsfils-spring-segment-routing-mpls-00

Abstract

Segment Routing (SR) leverages the source routing paradigm. A node steers a packet through a controlled set of instructions, called segments, by prepending the packet with an SR header. A segment can represent any instruction, topological or service-based. 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.

The Segment Routing architecture can be directly applied to the MPLS data plane with no change in the forwarding plane. This drafts describes how Segment Routing operates on top of the MPLS data plane.

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].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

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

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 23, 2014.

Copyright Notice

Table of Contents

1. Introduction

The Segment Routing architecture [I-D.filsfils-rtgwg-segment-routing] can be directly applied to the MPLS data plane with no change in the forwarding plane. This drafts describes how Segment Routing operates on top of the MPLS data plane.

The Segment Routing use cases are described in in [I-D.filsfils-rtgwg-segment-routing-use-cases].

Link State protocol extensions for Segment Routing are described in [I-D.previdi-isis-segment-routing-extensions], [I-D.psenak-ospf-segment-routing-extensions] and [I-D.psenak-ospf-segment-routing-ospfv3-extension].

1.1. Illustration

Segment Routing, applied to the MPLS data plane, offers the ability to tunnel services (VPN, VPLS, VPWS) from an ingress PE to an egress PE, without any other protocol than ISIS or OSPF ([I-D.previdi-isis-segment-routing-extensions] and [I-D.psenak-ospf-segment-routing-extensions]). LDP and RSVP-TE signaling protocols are not required.

Note that [draft-filsfils-rtgwg-segment-routing-ldp-interop-00] documents SR co-existence and interworking with other MPLS signaling protocols, if present in the network during a migration, or in case of non-homogeneous deployments.

The operator only needs to allocate one node segment per PE and the SR IGP control-plane automatically builds the required MPLS forwarding constructs from any PE to any PE.

```
                              P1---P2
                             /       \
             A---CE1---PE1             PE2---CE2---Z
                             \       /
                              P4---P4
```

Figure 1: IGP-based MPLS Tunneling

In Figure 1 above, the four nodes A, CE1, CE2 and Z are part of the same VPN.

PE2 advertises (in the IGP) a host address 192.0.2.2/32 with its attached node segment 102.

CE2 advertises to PE2 a route to Z. PE2 binds a local label LZ to that route and propagates the route and its label via MPBGP to PE1

with nhop 192.0.2.2 (PE2 loopback address).

PE1 installs the VPN prefix Z in the appropriate VRF and resolves the next-hop onto the node segment 102. Upon receiving a packet from A destined to Z, PE1 pushes two labels onto the packet: the top label is 102, the bottom label is LZ. 102 identifies the node segment to PE2 and hence transports the packet along the ECMP-aware shortest-path to PE2. PE2 then processes the VPN label LZ and forwards the packet to CE2.

Supporting MPLS services (VPN, VPLS, VPWS) with SR has the following benefits:

Simple operation: one single intra-domain protocol to operate: the IGP. No need to support IGP synchronization extensions as described in [RFC5443] and [RFC6138].

Excellent scaling: one Node-SID per PE.

2. MPLS Instantiation of Segment Routing

When applied to MPLS, the 20 right-most bits of the segment are encoded as a label. This implies that, in the MPLS instantiation, the SID values are allocated within a reduced 20-bit space out of the 32-bit SID space.

The notion of indexed global segment fits the MPLS architecture [RFC3031] as the absolute value allocated to any segment (global or local) can be managed by a local allocation process (similarly to other MPLS signaling protocols).

As described in [RFC3031] labels can be signaled by various protocols. Within a SR domain, LDP and RSVP MPLS signaling protocols are not required. If present, SR can coexist and interwork with LDP and RSVP [draft-filsfils-rtgwg-segment-routing-ldp-interop-00].

The source routing model described in [I-D.filsfils-rtgwg-segment-routing] is inherited from the ones proposed by [RFC1940] and [RFC2460]. The source routing model offers the support for explicit routing capability.

Contrary to RSVP-based explicit routes where tunnel midpoints maintain states, SR-based explicit routes only require per-flow states at the ingress edge router where the traffic engineer policy is applied.

Contrary to RSVP-based explicit routes which consist in non-ECMP

circuits (similar to ATM/FR), SR-based explicit routes can be built as list of ECMP-aware node segments and hence ECMP-aware traffic engineering is natively supported by SR.

When Segment Routing is instantiated over the MPLS data plane the following applies:

A list of segments is represented as a stack of labels.

The active segment is the top label.

The CONTINUE operation is implemented as an MPLS swap operation. When the same SRGB block is used throughout the SR domain, the outgoing label value is equal to the incoming label value . Else, the outgoing label value is [SRGB(next_hop)+index]

The NEXT operation is implemented as an MPLS pop operation.

The PUSH operation is implemented as an MPLS push of a label stack.

In conclusion, there are no changes in the operations of the data-plane currently used in MPLS networks.

3. Segment List History

In the abstract SR routing model [I-D.filsfils-rtgwg-segment-routing], any node N along the journey of the packet is able to determine where the packet P entered the SR domain and where it will exit. The intermediate node is also able to determine the paths from the ingress edge router to itself, and from itself to the egress edge router.

In the MPLS instantiation, as the packet travels through the SR domain, the stack is depleted and the segment list history is gradually lost.

Future version of this document will describe how this information can be preserved in MPLS domains.

4. IANA Considerations

TBD

5. Manageability Considerations

TBD

6. Security Considerations

TBD

7. Acknowledgements

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

8.2. Informative References

[I-D.filsfils-rtgwg-segment-routing]
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 Architecture", draft-filsfils-rtgwg-segment-routing-00 (work in progress), June 2013.

[I-D.filsfils-rtgwg-segment-routing-use-cases]
Filsfils, C., Francois, P., Previdi, S., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing Use Cases", draft-filsfils-rtgwg-segment-routing-use-cases-01 (work in progress), July 2013.

[I-D.previdi-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., and S. Litkowski, "IS-IS Extensions for Segment Routing", draft-previdi-isis-segment-routing-extensions-02 (work in

progress), July 2013.

[I-D.psenak-ospf-segment-routing-extensions]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., and W. Henderickx, "OSPF Extensions for Segment Routing", draft-psenak-ospf-segment-routing-extensions-03 (work in progress), October 2013.

[I-D.psenak-ospf-segment-routing-ospfv3-extension]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., and W. Henderickx, "OSPFv3 Extensions for Segment Routing", draft-psenak-ospf-segment-routing-ospfv3-extension-00 (work in progress), October 2013.

[RFC1940] Estrin, D., Li, T., Rekhter, Y., Varadhan, K., and D. Zappala, "Source Demand Routing: Packet Format and Forwarding Specification (Version 1)", RFC 1940, May 1996.

[RFC5443] Jork, M., Atlas, A., and L. Fang, "LDP IGP Synchronization", RFC 5443, March 2009.

[RFC6138] Kini, S. and W. Lu, "LDP IGP Synchronization for Broadcast Networks", RFC 6138, February 2011.

[draft-filsfils-rtgwg-segment-routing-ldp-interop-00]
Filsfils, C. and S. Previdi, "Segment Routing interoperability with LDP", October 2013.

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

Ahmed Bashandy
Cisco Systems, Inc.
170, West Tasman Drive
San Jose, CA 95134
US

Email: bashandy@cisco.com

Bruno Decraene
Orange
FR

Email: bruno.decraene@orange.com

Stephane Litkowski
Orange
FR

Email: stephane.litkowski@orange.com

Martin Horneffer
Deutsche Telekom
Hammer Str. 216-226
Muenster 48153
DE

Email: Martin.Horneffer@telekom.de

Igor Milojevic
Telekom Srbija
Takovska 2
Belgrade
RS

Email: igormilojevic@telekom.rs

Rob Shakir
British Telecom
London
UK

Email: rob.shakir@bt.com

Saku Ytti
TDC Oy
Mechelininkatu 1a
TDC 00094
FI

Email: saku@ytti.fi

Wim Henderickx
Alcatel-Lucent
Copernicuslaan 50
Antwerp 2018
BE

Email: wim.henderickx@alcatel-lucent.com

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134

US

Email: Jeff.Tantsura@ericsson.com

Edward Crabbe
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
US

Email: edc@google.com