Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 15, 2017

X. Xu, Ed.
S. Bryant, Ed.
Huawei
R. Raszuk
Bloomberg LP
U. Chunduri
Huawei
L. Contreras
Telefonica I+D
L. Jalil
Verizon
H. Assarpour
Broadcom
V. Gunter
Nokia
J. Tantsura
Individual
S. Ma
Juniper
June 13, 2017

# Unified Source Routing Instruction using MPLS Label Stack
## draft-xu-mpls-unified-source-routing-instruction-01

Abstract

MPLS-SPRING is an MPLS data plane-based source routing paradigm in which a sender of a packet is allowed to partially or completely specify the route the packet takes through the network by imposing stacked MPLS labels to the packet. MPLS-SPRING could be leveraged to realize a unified source routing mechanism across MPLS, IPv4 and IPv6 data planes by using a unified source routing instruction set while preserving backward compatibility with MPLS-SPRING.

Table of Contents



## 1. Introduction

MPLS-SPRING [I-D.ietf-spring-segment-routing-mpls] is an MPLS data plane-based source routing paradigm in which a sender of a packet is allowed to partially or completely specify the route the packet takes through the network by imposing stacked MPLS labels to the packet. MPLS-SPRING could be leveraged to realize a unified source routing mechanism across MPLS, IPv4 and IPv6 data planes by using a unified source routing instruction set while preserving backward compatibility with MPLS-SPRING. More specifically, the source routing instruction set information contained in a source routed packet could be uniformly encoded as an MPLS label stack no matter the underlay is IPv4, IPv6 or MPLS.

The traditional IPv4 and IPv6 source routing mechanisms by use of IPv4 Source Routing Options and IPv6 Route Header Type 0 Extension respectively have been deprecated due to their obvious security vulnerabilities. IPv6 SPRING (a.k.a., SRv6) [I-D.ietf-6man-segment-routing-header] is a newly proposed IPv6 source routing mechanism in which the source route instruction information is encoded as an ordered list of 128-bit long IPv6 addresses and contained in the Source Routing Header (SRH). Although it has overcome the security vulnerability issues associated with the traditional IPv6 source routing mechanism as claimed in [I-D.ietf-6man-segment-routing-header], it still has the following obvious drawbacks which need to be addressed: 1) the encapsulation overhead is significant especially when the list of the explicit routing hops is very long; 2) for those transit IPv6 routers that don't support the flow label-based load-balancing mechanism yet, the ECMP load-balancing effect may be impacted seriously if they could not recognize the SRH and therefore could not obtain the five tuple of the source routed IPv6 packet; 3) it requires a totally new forwarding logic on basis of the SRH and the forwarding performance associated with the IPv6 SRH may still be a big concern for some hardware platforms.

Section 3 describes various use cases for the unified source routing instruction mechanism and Section 4 describes a typical application scenario and how the packet forwarding happens.

### 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. Terminology

This memo makes use of the terms defined in [RFC3031] and [I-D.ietf-spring-segment-routing-mpls].

## 3. Use Cases

The unified source routing mechanism across IPv4, IPv6 and MPLS is useful at least in the following use cases:

- o Incremental deployment of the MPLS-SPRING technology. Since there is no need to run any other label distribution protocol (e.g., LDP, see [I-D.ietf-spring-segment-routing-ldp-interop] for more details.) on those non-MPLS-SPRING routers for incremental deployment purposes, the network provisioning is greatly simplified, which is one of the major claimed benefits of the

MPLS-SPRING technology (i.e., running a single protocol). In fact, this unified source routing mechanism is even useful in a fully upgraded MPLS-SPRING network since the headache associated with the MPLS-SPRING load-balancing as described in [I-D.ietf-mpls-spring-entropy-label] can now be avoided by using the source port of the UDP tunnel header as an entropy field instead.

- A poor man's light-weight alternative to SRv6 [I-D.ietf-6man-segment-routing-header]. At least, it could be deployed as an interim until full featured SRv6 is available on more platforms. Since the Source Routing Header (SRH) [I-D.ietf-6man-segment-routing-header] consisting of an ordered list of 128-bit long IPv6 addresses is now replaced by an ordered list of 32-bit long label entries (i.e., label stack), the encapsulation overhead and forwarding performance issues associated with SRv6 are eliminated.

- A new IPv4 source routing mechanism which has overcome the security vulnerability issues associated with the traditional IPv4 source routing mechanism.

- Traffic Engineering scenarios where only a few routers (e.g., the entry and exit nodes of each plane in the dual-plane network ) are specified as segments of explicit paths. In this way, only a few routers are required to support the MPLS-SPRING capability while all the other routers just need to support IP forwarding capability, which would significantly reduce the deployment cost of this new technology.

- MPLS-based Service Function Chaining (SFC) [I-D.xu-mpls-service-chaining]. Based on the unified source routing mechanism as described in this document, only SFC-related nodes including Service Function Forwarders (SFF), Service Functions (SF) and classifiers are required to recognize the SFC encapsulation header in the MPLS label stack form, while the intermediate routers just need to support vanilla IP forwarding (either IPv4 or IPv6). In other words, it undoubtedly complies with the transport-independence requirement as listed in the SFC architecture document [RFC7665].

## 4. Packet Forwarding Procedures

```
     +-----+       +-----+       +-----+        +-----+        +-----+
     |  A  +-------+  B  +-------+  C  +--------+  D  +--------+  H  |
     +-----+       +--+--+       +--+--+        +--+--+        +-----+
                      |             |              |
                      |             |              |
                   +--+--+       +--+--+        +--+--+
                   |  E  +-------+  F  +--------+  G  |
                   +-----+       +-----+        +-----+

          +--------+
          |IP(A->E)|
          +--------+                 +--------+
          |  L(G)  |                 |IP(E->G)|
          +--------+                 +--------+        +--------+
          |  L(H)  |                 |  L(H)  |        |IP(G->H)|
          +--------+                 +--------+        +--------+
          | Packet |     --->        | Packet |  --->  | Packet |
          +--------+                 +--------+        +--------+
```

Figure 1

As shown in Figure 1, Assume Router A, E, G and H are MPLS-SPRING-capable routers while the remaining are only capable of forwarding IP packets. Router A, E, G and H advertise their Segment Routing related information via IS-IS or OSPF. Now assume router A wants to send a given IP or MPLS packet via an explicit path of {E->G->H}, router A would impose an MPLS label stack corresponding to that explicit path on the received IP packet. Since there is no Label Switching Path (LSP) towards router E, router A would replace the top label indicating router E with an IP-based tunnel for MPLS (e.g., MPLS-over-UDP [RFC7510] or MPLS-over-GRE [RFC4023]) towards router E and then send it out. In other words, router A would pop the top label and then encapsulate the MPLS packet with an IP-based tunnel towards router E. When the IP-encapsulated MPLS packet arrives at router E, router E would strip the IP-based tunnel header and then process the decapsulated MPLS packet accordingly. Since there is no LSP towards router G which is indicated by the current top label of the decapsulated MPLS packet, router E would replace the current top label with an IP-based tunnel towards router G and send it out. When the packet arrives at router G, router G would strip the IP-based tunnel header and then process the decapsulated MPLS packet. Since there is no LSP towards router H, router G would replace the current top label with an IP-based tunnel towards router H. Now the packet encapsulated with the IP-based tunnel towards router H is exactly the original packet that router A had intended to send towards router H. If the packet is an MPLS packet, router G could use any IP-based tunnel for MPLS (e.g., MPLS-over-UDP [RFC7510] or MPLS-over-GRE [RFC4023]). If the packet is an IP packet, router G could use any IP tunnel for IP (e.g., IP-in-UDP [I-D.xu-intarea-ip-in-udp] or GRE

[RFC2784]).  That original IP or MPLS packet would be forwarded towards router H via an IP-based tunnel.  When the encapsulated packet arrives at router H, router H would decapsulate it into the original packet and then process it accordingly.

Note that in the above description, it's assumed that the label associated with each prefix-SID advertised by the owner of the prefix-SID is a Penultimate Hop Popping (PHP) label (e.g., the NP-flag [I-D.ietf-ospf-segment-routing-extensions] associated with the corresponding prefix SID is not set).  Figure 2 demostrates the packet walk in the case where the label associated with each prefix-SID advertised by the owner of the prefix-SID is not a Penultimate Hop Popping (PHP) label (e.g., the NP-flag [I-D.ietf-ospf-segment-routing-extensions] associated with the corresponding prefix SID is set).  Although the above description is based on the use of prefix-SIDs, the unified source routing instruction approach is actually applicable to the use of adj-SIDs as well.  For instance, when the top label of a received MPLS packet indicates an given adj-SID and the corresponding adjacent node to that adj-SID is not MPLS-capable, the top label would be replaced by an IP-based tunnel towards that adjacent node and then forwarded over the correponding link indicated by that adj-SID.

```
  +-----+       +-----+       +-----+        +-----+        +-----+
  |  A  +-------+  B  +-------+  C  +--------+  D  +--------+  H  |
  +-----+       +--+--+       +--+--+        +--+--+        +-----+
                   |             |              |
                   |             |              |
                +--+--+       +--+--+        +--+--+
                |  E  +-------+  F  +--------+  G  |
                +-----+       +-----+        +-----+

       +--------+
       |IP(A->E)|
       +--------+                  +--------+
       |  L(E)  |                  |IP(E->G)|
       +--------+                  +--------+        +--------+
       |  L(G)  |                  |  L(G)  |        |IP(G->H)|
       +--------+                  +--------+        +--------+
       |  L(H)  |                  |  L(H)  |        |  L(H)  |
       +--------+                  +--------+        +--------+
       | Packet |     --->         | Packet |  --->  | Packet |
       +--------+                  +--------+        +--------+
```

Figure 2

Note that as for which tunnel encapsulation type should be used, it could be manually specified on tunnel ingress routers or be learnt from the tunnel egress routers' advertisements of its tunnel

encapsulation capability. How to advertise the tunnel encapsulation capability using IS-IS or OSPF are specified in [I-D.ietf-isis-encapsulation-cap] and [I-D.ietf-ospf-encapsulation-cap] respectively.

## 5. Acknowledgements

Thanks Joel Halpern, Bruno Decraene and Loa Andersson for their insightful comments on this draft.

## 6. IANA Considerations

No IANA action is required.

## 7. Security Considerations

TBD.

## 8. References

### 8.1. 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>.

### 8.2. Informative References

[I-D.ietf-6man-segment-routing-header]
Previdi, S., Filsfils, C., Raza, K., Leddy, J., Field, B., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Matsushima, S., Leung, I., Linkova, J., Aries, E., Kosugi, T., Vyncke, E., Lebrun, D., Steinberg, D., and R. Raszuk, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-06 (work in progress), March 2017.

[I-D.ietf-isis-encapsulation-cap]
Xu, X., Decraene, B., Raszuk, R., Chunduri, U., Contreras, L., and L. Jalil, "Advertising Tunnelling Capability in IS-IS", draft-ietf-isis-encapsulation-cap-01 (work in progress), April 2017.

[I-D.ietf-mpls-spring-entropy-label]
Kini, S., Kompella, K., Sivabalan, S., Litkowski, S., Shakir, R., and j. jefftant@gmail.com, "Entropy label for SPRING tunnels", draft-ietf-mpls-spring-entropy-label-06 (work in progress), May 2017.

[I-D.ietf-ospf-encapsulation-cap]
Xu, X., Decraene, B., Raszuk, R., Contreras, L., and L. Jalil, "Advertising Tunneling Capability in OSPF", draft-ietf-ospf-encapsulation-cap-03 (work in progress), May 2017.

[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-16 (work in progress), May 2017.

[I-D.ietf-spring-segment-routing-ldp-interop]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., and S. Litkowski, "Segment Routing interworking with LDP", draft-ietf-spring-segment-routing-ldp-interop-07 (work in progress), May 2017.

[I-D.ietf-spring-segment-routing-mpls]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls-08 (work in progress), March 2017.

[I-D.xu-intarea-ip-in-udp]
Xu, X., Lee, Y., and F. Yongbing, "Encapsulating IP in UDP", draft-xu-intarea-ip-in-udp-04 (work in progress), December 2016.

[I-D.xu-mpls-service-chaining]
Xu, X., Bryant, S., Assarpour, H., Shah, H., Contreras, L., daniel.bernier@bell.ca, d., jefftant@gmail.com, j., and S. Ma, "Service Chaining using an Unified Source Routing Instruction", draft-xu-mpls-service-chaining-02 (work in progress), May 2017.

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, DOI 10.17487/RFC2784, March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <http://www.rfc-editor.org/info/rfc3031>.

[RFC4023] Worster, T., Rekhter, Y., and E. Rosen, Ed., "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)", RFC 4023, DOI 10.17487/RFC4023, March 2005, <http://www.rfc-editor.org/info/rfc4023>.

[RFC4817] Townsley, M., Pignataro, C., Wainner, S., Seely, T., and J. Young, "Encapsulation of MPLS over Layer 2 Tunneling Protocol Version 3", RFC 4817, DOI 10.17487/RFC4817, March 2007, <http://www.rfc-editor.org/info/rfc4817>.

[RFC7510] Xu, X., Sheth, N., Yong, L., Callon, R., and D. Black, "Encapsulating MPLS in UDP", RFC 7510, DOI 10.17487/RFC7510, April 2015, <http://www.rfc-editor.org/info/rfc7510>.

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015, <http://www.rfc-editor.org/info/rfc7665>.

Authors' Addresses

Xiaohu Xu (editor)
Huawei

Email: xuxiaohu@huawei.com

Stewart Bryant (editor)
Huawei

Email: stewart.bryant@gmail.com

Robert Raszuk
Bloomberg LP

Email: robert@raszuk.net

Uma Chunduri
Huawei

Email: uma.chunduri@gmail.com

Luis M. Contreras
Telefonica I+D

Email: luismiguel.contrerasmurillo@telefonica.com

Luay Jalil
Verizon

Email: luay.jalil@verizon.com

Hamid Assarpour
Broadcom

Email: hamid.assarpour@broadcom.com

Van De Velde, Gunter
Nokia

Email: gunter.van_de_velde@nokia.com

Jeff Tantsura
Individual

Email: jefftant.ietf@gmail.com

Shaowen Ma
Juniper

Email: mashao@juniper.net