SPRING
Internet-Draft
Intended status: Standards Track
Expires: January 9, 2020

C. Filsfils, Ed.
Cisco Systems, Inc.
D. Cai
Alibaba
Z. Jiang
Tencent
D. Voyer
Bell Canada
A. Shawky
Saudi Telecom Company
N. Leymann
Deutsche Telekom
D. Steinberg
Lapishills Consulting Limited
S. Zandi
G. Dawra
LinkedIn
I. Meilik
Broadcom
J. Uttaro
AT&T
L. Jalil
Verizon
N. So
Reliance
M. Fiumano
Sprint
M. Khaddam
Cox
J. Ma
China Unicom
S. Matsushima
Softbank
F. Ferguson
CenturyLink
T. Miyasaka
KDDI
K. Ebisawa
Toyota Motor Corporation
Y. Ueno
NTT Communications Corporation
W. Henderickx
Nokia
P. Jonnalagadda
Barefoot Networks
J. Bhattacharya
K. Raza
P. Camarillo, Ed.
Cisco Systems, Inc.
July 8, 2019

Abstract

The SRv6 "micro segment" (SRv6 uSID or uSID for short) instruction is defined and illustrated.

It is a straightforward extension to the SRv6 Network Programming model and its SRH encapsulation.

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.

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 https://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 January 9, 2020.

Copyright Notice

Table of Contents

1. Introduction

SRv6 Network Programming [I-D.ietf-spring-srv6-network-programming] defines a mechanism to build a network program with topological and service segments. It leverages the SRH [I-D.ietf-6man-segment-routing-header] to encode a network program together with optional metadata shared among the different SIDs.

This draft extends SRv6 Network Programming with a new type of SRv6 SID behavior: SRv6 uN. This is combined with the rest of instructions of the network program and the SRH encapsulation to build programs in a scalable and efficient way.

2. Terminology

The SRv6 Network Programming [I-D.ietf-spring-srv6-network-programming] and SRH [I-D.ietf-6man-segment-routing-header] terminology is leveraged and extended with the following terms:

uSID carrier: a 128bit SRv6 SID of format <uSID-Block><Active-uSID><Next-uSID>...<Last-uSID><End-of-Carrier>...<End-of-Carrier>.

uSID block: A block of uSID's

It can be any IPv6 prefix allocated to the provider (e.g. /40 or /48), or it can be any block generally available for private use. An SR domain may have multiple uSID blocks.

In this document we leverage FC00::/8 block reserved for private use as ULA space (RFC4193). Throughout this document we use FC00::/16 as the illustrated uSID block. ULA space allows for up to 256 uSID blocks in FC00::/8.

uSID: in this document a 16-bit ID. A different length may be used.

Active uSID: first uSID after the uSID block

Next uSID: next uSID after the Active uSID

Last uSID: from left to right, the last uSID before the first End-of-Carrier uSID

End-of-Carrier: reserved ID used to mark the end of a uSID carrier. The value 0000 is selected as End-of-Carrier. All of the empty uSID carrier positions must be filled with the End-of-Carrier ID. Hence, the End-of-Carrier can be present more than once in a uSID carrier.

Parent (node): the node at which an uSID is instantiated. The uSIDs are instantiated on a per-parent node basis.

Behavior of an uSID: the SRv6 function associated with a given ID. Section 3 defines them.

2.1. Notation for human readability

For human readability, the example in this document follow this notation:

FC00::/16 is the uSID block used in the SR domain

0N00: uN behavior bound to node N

3. SRv6 behaviors associated with a uSID

The SRv6 SRH encapsulation and its network programming model are extended with the following functions:

3.1. uN

The uN behavior is a variant of the endpoint behavior.

This behavior takes a 96b argument, "Arg", which contains the next uSIDs in the uSID carrier.

When N receives a packet whose IPv6 DA is S and S is a local uN SID, N does:

```
1.   IF DA[32..47] != 0                                  ;; Ref1
2.      Copy DA[32..127] into DA[16..111]
3.      Set DA[112..127] to 0x0000
4.      Forward the packet to the new DA
5.   ELSE
6.      Execute the End pseudocode                       ;; Ref2
```

Ref 1: DA[X..Y] refers to the bits from position X to Y (included) in the IPv6 Destination Address of the received packet. The bit 0 is the MSB, while the bit 127 is the LSB.

Ref 2: This refers to the End behavior as defined in Section 4.1 of [I-D.ietf-spring-srv6-network-programming]. The End behavior may be combined with the PSP, USP and USD flavours.

4. Routing

If N is configured with a uN SID FC00:0N00::/32 then the operator must ensure that N advertises FC00:0N00::/32 in routing.

5. Illustration

This section extends the illustrations for SRv6 Network Programming [I-D.filsfils-spring-srv6-net-pgm-illustration] to cover uSID. The reference topology is the same with the addition of link 6-8.

5.1. Reference diagram

Nodes 1 to 8 are considered within the network domain.

Nodes X and Y are outside the domain.

Nodes 1 and 8 act as PE respectively to nodes X and Y.

All the links within the domain have the same IGP metric. The IGP-metric shortest-path from 1 to 8 is 1-2-7-8 while the latency-metric shortest-path from 1 to 8 is 1-2-3-4-5-6-7-8.

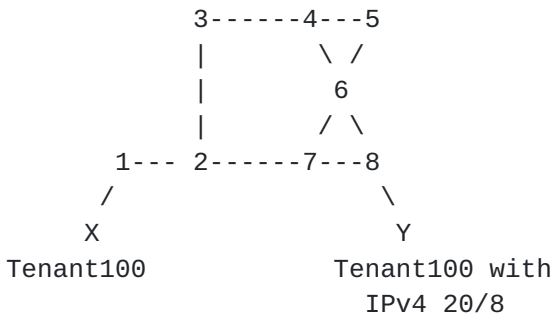

Figure 1: Reference topology

5.2. SRv6 overlay with underlay optimization

Let us illustrate a low-latency SR-L3VPN service delivered to a packet (X,Y).

PE 1 encapsulates (X, Y) in an outer IPv6 header with DA = FC00:0300:0500:0700:: and SRH (B:8:D0::; SL=1; NH=4). Leveraging the illustration conventions from SRv6 network programming, the following resulting packet leaves node 1 in the direction of node 3:

(A1::, FC00:0300:0500:0700::)(B:8:D0::; SL=1; NH=4)(X, Y)

FC00:0300:0500:0700:: is a uSID carrier encoding a source routed stateless path via node 3 then 5 then 7.

B:8:D0:: is an End.DT4 SID instantiated at node 8.

1 sends this packet to 2, as 2 is on the shortest-path to FC00:0300::/32 advertised by 3.

When 2 receives the packet, 2 performs a regular IPv6 FIB lookup. It finds a FIB entry for FC00:0300::/32 and forwards along the shortest path to 3.

When 3 receives the packet, 3 matches FC00:0300::/32 in its "My SID Table" and executes the uN behavior. The updated DA becomes FC00:0500:0700::. Node 3 then performs a lookup on the updated DA and forwards the packet to 5 along the shortest path to FC00:0500::/32.

The following packet leaves node 3:

(A1::, FC00:0500:0700::)(B:8:D0::; SL=1; NH=4)(X, Y)

4 forwards along the shortest path to FC00:0500::/32.

When 5 receives the packet, 5 matches FC00:0500::/32 in its "My SID Table" and executes the uN behavior. The updated DA becomes FC00:0700::. 5 performs a lookup on the updated DA and forwards the packet to 7 along the shortest path to FC00:0700::/32.

The following packet leaves node 5:

(A1::, FC00:0700::)(B:8:D0::; SL=1; NH=4)(X, Y)

6 forwards along the shortest path to FC00:0700::/32.

When 7 receives the packet, 7 matches FC00:0700::/32 in its "My SID Table" and finds the bound function uN. As a result, Node 7 executes the "End with PSP and USD support" pseudocode, decrementing the SL value in the SRH, and updating the DA with the next SID B:8:D0::. Since the SL value is zero the SRH is removed. Node 7 performs a lookup on the updated DA and forwards along the shortest path.

The following packet leaves node 7:

(A1::, B:8:D0::)(X, Y)

8 receives it, performs the End.DT4 function and sends the IP packet (X, Y) towards its VPN destination.

This example illustrates the benefits highlighted in the next section.

6. Benefits

Perfect integration with SRv6 Network Programming

SRv6 uSID is an instruction of the SRv6 network programming model

Perfect integration with SRH

Any SID in DA or SRH can be an SRv6 uSID carrier

Scalable SR Policy

7 uSID' per uSID carrier

21 source routing waypoints in solely 40bytes of overhead

T.Encaps.Red with an SRH of 40 bytes (8 fixed + 2 * 16 bytes)

7 uSID's in DA and 14 in SRH

Efficient MTU overhead

In apple to apple comparison, the SRv6 solution outperforms any alternative (VxLAN with SR-MPLS, CRH).

Scalable number of globally unique nodes in the domain

16-bit uSID: 65k uSIDs per domain block (*256 solely using FC/8)

32-bit uSID: 4.3M uSIDs per domain block (*256 solely using FC/8)

Hardware-friendly:

Leverages mature hardware capabilities (shift)

Avoids any extra lookup in indexed mapping table

Demonstrated by Cisco linerate implementation on Jericho1

Control-Plane friendly

No indexed mapping table is required

No routing extension is required: a simple /32 advertisement suffices

7. Security

The security rules defined in Section 7 of [I-D.ietf-spring-srv6-network-programming], protect intra-domain deployments that includes SRv6 uSID.

8. Acknowledgements

The authors would like to acknowledge Francois Clad, Peter Psenak, Ketan Talaulikar, Swadesh Agrawal, Zafar Ali, Darren Dukes, Kiran Sadshiran, Junaid Israr, Lakshmanan Srikanth, Asif Islam, Saleem Hafeez, Michael MacKenzie, Sushek Shekar, YuanChao Su, Alexander Preusche, Alberto Donzelli, Miya Kohno, David Smith, Ianik Semco, Bertrand Duvivier, Frederic Trate, Kris Michielsen, Eyal Dagan, Eli Stein, Ofer Iny, Elad Naor, Aviad Behar, Joseph Chin.

9. Contributors

Tomonobu Niwa
KDDI
Japan

Email: to-niwa@kddi.com

10. References

10.1. Normative References

[I-D.ietf-6man-segment-routing-header]
Filsfils, C., Dukes, D., Previdi, S., Leddy, J., Matsushima, S., and d. daniel.voyer@bell.ca, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-21 (work in progress), June 2019.

[I-D.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., daniel.voyer@bell.ca, d., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-ietf-spring-srv6-network-programming-01 (work in progress), July 2019.

[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>.

[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>.

10.2. Informative References

[I-D.filsfils-spring-srv6-net-pgm-illustration]
Filsfils, C., Camarillo, P., Li, Z., Matsushima, S., Decraene, B., Steinberg, D., Lebrun, D., Raszuk, R., and J. Leddy, "Illustrations for SRv6 Network Programming", draft-filsfils-spring-srv6-net-pgm-illustration-00 (work in progress), February 2019.

Authors' Addresses

Clarence Filsfils (editor)
Cisco Systems, Inc.
Belgium

Email: cf@cisco.com

Dennis Cai
Alibaba
China

Email: d.cai@alibaba-inc.com

Zhichun Jiang
Tencent
China

Email: zcjiang@tencent.com

Daniel Voyer
Bell Canada
Canada

Email: daniel.voyer@bell.ca

Ahmed Shawky
Saudi Telecom Company
Saudi Arabia

Email: ashawky@stc.com.sa

Nic Leymann
Deutsche Telekom
Germany

Email: N.Leymann@telekom.de

Dirk Steinberg
Lapishills Consulting Limited
Cyprus

Email: dirk@lapishills.com

Shawn Zandi
LinkedIn
United States of America

Email: szandi@linkedin.com

Gaurav Dawra
LinkedIn
United States of America

Email: gdawra@linkedin.com

Israel Meilik
Broadcom
Israel

Email: israel.meilik@broadcom.com

Jim Uttaro
AT&T
United States of America

Email: ju1738@att.com

Luay Jalil
Verizon
United States of America

Email: luay.jalil@one.verizon.com

Ning So
Reliance
United States of America

Email: Ning.So@ril.com

Michael Fiumano
Sprint
United States of America

Email: michael.f.fiumano@sprint.com

Mazen Khaddam
Cox
United States of America

Email: Mazen.Khaddam@cox.com

Jichun Ma
China Unicom
China

Email: majc16@chinaunicom.cn

Satoru Matsushima
Softbank
Japan

Email: satoru.matsushima@g.softbank.co.jp

Francis Ferguson
CenturyLink
United States of America

Email: Francis.Ferguson@centurylink.com

Takuya Miyasaka
KDDI
Japan

Email: ta-miyasaka@kddi.com

Kentaro Ebisawa
Toyota Motor Corporation
Japan

Email: ebisawa@toyota-tokyo.tech

Yukito Ueno
NTT Communications Corporation
Japan

Email: yukito.ueno@ntt.com

Wim Henderickx
Nokia
Belgium

Email: wim.henderickx@nokia.com

Prem Jonnalagadda
Barefoot Networks
United States of America

Email: prem@barefootnetworks.com

Jisu Bhattacharya
Cisco Systems, Inc.
United States of America

Email: jisu@cisco.com

Kamran Raza
Cisco Systems, Inc.
Canada

Email: skraza@cisco.com

Pablo Camarillo (editor)
Cisco Systems, Inc.
Spain

Email: pcamaril@cisco.com