SPRING
Internet-Draft
Intended status: Standards Track
Expires: August 28, 2020

C. Filsfils, Ed.
P. Camarillo, Ed.
Cisco Systems, Inc.
D. Cai
Alibaba
D. Voyer
Bell Canada
I. Meilik
Broadcom
K. Patel
Arrcus, Inc.
W. Henderickx
Nokia
P. Jonnalagadda
Barefoot Networks
D. Melman
Marvell

February 25, 2020

Network Programming extension: SRv6 uSID instruction
draft-filsfils-spring-net-pgm-extension-srv6-usid-04

Abstract

The SRv6 "micro segment" (SRv6 uSID or uSID for short) instruction is a straightforward extension of the SRv6 Network Programming model:

- The SRv6 Control Plane is leveraged without any change
- The SRH dataplane encapsulation is leveraged without any change
- Any SID in the SID list can carry micro segments

This enables:

- ultra-scale (e.g. multi-domain 5G deployments)
- minimum MTU overhead
- installed-base reuse

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 August 28, 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 extension fully leverages the SRv6 network programming solution:

- The SRv6 Control Plane is leveraged without any change
- The SRH dataplane encapsulation is leveraged without any change
- Any SID in the SID list can carry micro segments

This enables:

- ultra-scale (e.g. multi-domain 5G deployments)
- minimum MTU overhead
- installed-base reuse

2. Terminology

The SRv6 Network Programming and SRH terminology is leveraged and extended with the following terms:

Term	Definition
uSID block	A block of uSID's. It can be any IPv6 prefix available to the provider. In this note we will assume a /32 sub-allocated from a public block [I-D.matsushima-spring-srv6-deployment-status]. Other block length could be used.
uSID	In this document a 16-bit ID. A different uSID 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 uSID 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.
uSID carrier	A 128bit SRv6 SID of format <uSID-Block><Active-uSID><Next-uSID>...<Last-uSID><End-of-Carrier>...<End-of-Carrier>. A uSID carrier can be encoded in the Destination Address of an IPv6 header or at any position in the Segment List of an SRH.

2.1. Notation for human readability

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

2001:db8::/32 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 80b argument, "Arg", which contains the next uSIDs in the uSID carrier.

When N receives a packet whose IPv6 DA matches a local uN SID, N does:

```
1.   IF DA[48..63] != 0                                    ;; Ref1
2.      Copy DA[48..127] into DA[32..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 Node 1 is configured with a uN SID 2001:db8:0100::/48 then the operator must ensure that Node 1 advertises 2001:db8:0100::/48 in routing.

5. Benefits

- Leverages SRv6 Network Programming with NO change
  - SRv6 uSID is an instruction of the SRv6 network programming model
- Leverages SRv6 dataplane (SRH) with NO change
  - Any SID in DA or SRH can be an SRv6 uSID carrier
- Leverages SRv6 Control-Plane with NO change
- Ultra-Scale
  - 6 uSID' per uSID carrier
  - 18 source routing waypoints in solely 40bytes of overhead
    - T.Encaps.Red with an SRH of 40 bytes (8 fixed + 2 * 16 bytes)
    - 6 uSID's in DA and 12 in SRH
- Lowest 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
  - 32-bit uSID: 4.3M uSIDs per domain block
- Proven Hardware-friendliness
  - Leverages mature hardware capabilities (shift, DA longest match)
  - Avoids any extra lookup in indexed mapping table
  - Demonstrated by the number of linerate interoperable hardware implementations at the first Interop report in February 2020, less than 9 months after the first public version of this document.
  - Public operator report of leverage of installed base

    * A micro-program which requires less than 6 uSID's only requires legacy IPinIP encapsulation behavior

- Scalable Control-Plane
    * No indexed mapping table is required
    * Summarization at area/domain boundary provides massive scaling advantage
    * No routing extension is required: a simple prefix advertisement suffices
- Seamless Deployment
    * A uSID may be used as a SID: i.e. the carrier holds a single uSID
    * The inner structure of an SR Policy can stay opaque to the source: i.e. a carrier with uSID's is just seen as a SID by the policy headend
- Security
    * Leverages SRv6's native SR domain security
- Large-Scale DC
    * SID's may be used to address applications on hosts (scale in 2^128)
    * Hardware friendliness of uSID's may be used to specify billions of waypoints in cost/power-optimized DC fabric

6. Running code

The hardware and software platforms listed below have demonstrated support for the uN instruction defined in this document.

Further on, all these implementations have participated in a joint interoperability testing.

Hardware implementations (in alphabetical order):

- Arrcus ArcOS (based on Broadcom Jericho2)
- Barefoot Tofino NPU

- Cisco 8000 Series Routers (based on Cisco Silicon One Q100)
- Cisco ASR9000 platform (with 3rd gen Tomahawk and 4th gen Lightspeed line-cards)
- Cisco NCS5500 platform (based on Broadcom Jericho/Jericho+)
- Marvell Prestera Falcon CX 8500

Software open-source implementations (in alphabetical order):

- FD.io VPP
- Linux Kernel

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. Work in progress

Future version of this document will include other uSID behaviors related to TE, VPN and service programming.

9. Acknowledgements

The authors would like to acknowledge Francois Clad, Peter Psenak, Ketan Talaulikar, Jakub Horn, 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.

10. Contributors

Jisu Bhattacharyaa
Cisco Systems, Inc.
United States of America

Email: jisu@cisco.com

Kamran Raza
Cisco Systems, Inc.

Canada

Email: skraza@cisco.com

John Bettink
Cisco Systems, Inc.
United States of America

Email: jbettink@cisco.com

Tomonobu Niwa
KDDI
Japan

Email: to-niwa@kddi.com

Luay Jalil
Verizon
United States of America

Email: luay.jalil@one.verizon.com

Zhichun Jiang
Tencent
China

Email: zcjiang@tencent.com

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

Jim Uttaro
AT&T
United States of America

Email: ju1738@att.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

11. References

11.1. Normative References

[I-D.ietf-6man-segment-routing-header]
Filsfils, C., Dukes, D., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-26 (work in progress), October 2019.

[I-D.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-ietf-spring-srv6-network-programming-10 (work in progress), February 2020.

[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>.

11.2. Informative References

[I-D.matsushima-spring-srv6-deployment-status]
Matsushima, S., Filsfils, C., Ali, Z., and Z. Li, "SRv6 Implementation and Deployment Status", draft-matsushima-spring-srv6-deployment-status-05 (work in progress), January 2020.

Authors' Addresses

Clarence Filsfils (editor)
Cisco Systems, Inc.
Belgium

Email: cf@cisco.com

Pablo Camarillo (editor)
Cisco Systems, Inc.
Spain

Email: pcamaril@cisco.com

Dennis Cai
Alibaba
China

Email: d.cai@alibaba-inc.com

Daniel Voyer
Bell Canada
Canada

Email: daniel.voyer@bell.ca

Israel Meilik
Broadcom
Israel

Email: israel.meilik@broadcom.com

Keyur Patel
Arrcus, Inc.
United States of America

Email: keyur@arrcus.com

Wim Henderickx
Nokia
Belgium

Email: wim.henderickx@nokia.com

Prem Jonnalagadda
Barefoot Networks
United States of America

Email: prem@barefootnetworks.com

David Melman
Marvell
Israel

Email: davidme@marvell.com