6man
Internet-Draft
Intended status: Standards Track
Expires: May 23, 2020

R. Bonica
Juniper Networks
Y. Kamite
NTT Communications Corporation
L. Jalil
C. Lenart
Verizon
N. So
F. Xu
Reliance Jio
G. Presbury
Hughes Network Systems
G. Chen
Baidu
Y. Zhu
China Telecom
Y. Zhou
ByteDance

November 20, 2019

The Per-Path Service Instruction (PPSI) Option
draft-bonica-6man-vpn-dest-opt-08

Abstract

SRm6 encodes Per-Path Service Instructions (PPSI) in a new IPv6 option, called the PPSI Option. This document describes the PPSI Option.

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 May 23, 2020.

Table of Contents

1. Introduction

An SRm6 [I-D.bonica-spring-srv6-plus] path provides unidirectional connectivity from its ingress node to its egress node. While an SRm6 path can follow the least cost path from ingress to egress, it can also follow any other path.

SRm6 paths are encoded as IPv6 [RFC8200] header chains. When an SRm6 ingress node receives a packet, it encapsulates the packet in an IPv6 header chain. It then forwards the encapsulated packet to the path's egress node. When the egress node receives the packet, it processes the SRm6 payload (i.e., the original packet).

SRm6 paths are programmable. They support several instruction types, including Per-Path Service Instructions (PPSI). PPSIs determine how

path egress nodes process SRm6 payloads. In the absence of a PPSI, the egress node processes SRm6 payloads as described in [RFC8200].

The following are examples of PPSIs:

- Remove any SRm6 encapsulation and forward the SRm6 payload through a specified interface.

- Remove any SRm6 encapsulation and forward the SRm6 payload using a specified routing table.

SRm6 encodes PPSIs in a new IPv6 option, called the PPSI Option. This document describes the PPSI Option.

PPSIs can be used to support Virtual Private Networks (VPN). Therefore, Appendix A of this document describes VPN technology and how PPSIs can be used to support a VPN.

2. 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.

3. PPSI Identifiers

PPSI Identifiers identify PPSIs. When a path egress node instantiates a PPSI, it also allocates a PPSI Identifier and associates the PPSI with the identifier.

PPSI Identifiers have node-local significance. This means that a path egress node must assign a unique PPSI Identifier to each PPSI that it instantiates. However, one path egress node can assign a PPSI Identifier to an instruction that it instantiates, while another path egress node can assign the same PPSI Identifier to a different PPSI that it instantiates.

4. The PPSI Option

The PPSI Option contains the following fields:

- Option Type: 8-bit selector. PPSI option. Value TBD by IANA. (Suggested value: 144). See Note below.

- Opt Data Len - 8-bit unsigned integer. Length of the option, in octets, excluding the Option Type and Option Length fields. This field MUST be set to 4.

- PPSI identifier - (32-bit selector). Identifies a PPSI.

The SRm6 PPSI option MAY appear in a Destination Options header that precedes an upper-layer header. It MUST NOT appear in a Hop-by-hop Options header or in a Destination Options header that precedes a Routing header.

When the SRm6 PPSI option appears in a Destination Options header, it MUST be the only option listed in the header. This is because the PPSI defines all path egress node behaviors.

NOTE : The highest-order two bits of the Option Type (i.e., the "act" bits) are 10. These bits specify the action taken by a destination node that does not recognize the option. The required action is to discard the packet and, regardless of whether or not the packet's Destination Address was a multicast address, send an ICMPv6 [RFC4443] Parameter Problem, Code 2, message to the packet's Source Address, pointing to the unrecognized Option Type.

The third highest-order bit of the Option Type (i.e., the "chg" bit) is 0. This indicates that Option Data cannot be modified along the path between the packet's source and its destination.

5. Destination Option Header Considerations

As per [RFC8200], the Destination Options header includes a Next Header field. The Next Header field identifies the header following the Destination Options header.

SRm6 can carry Ethernet payload after a Destination option header. Therefore, this document requests IANA to assign a protocol number for Ethernet. (The suggested value is 143.)

6. ICMPv6 Considerations

SRm6 implementations MUST comply with the ICMPv6 processing rules specified in Section 2.4 of [RFC4443]. For example:

- An SRm6 implementation MUST NOT originate an ICMPv6 error message in response to another ICMPv6 error message.

- An SRm6 implementation MUST rate limit the ICMPv6 messages that it originates.

7. Security Considerations

SRm6 domains MUST NOT span security domains. In order to enforce this requirement, security domain edge routers MUST do one of the following:

- Discard all inbound SRm6 packets
- Authenticate [RFC4302] [RFC4303] all inbound SRm6 packets

8. IANA Considerations

IANA is requested to allocate a code point from the Destination Options and Hop-by-hop Options registry (https://www.iana.org/assignments/ipv6-parameters/ipv6-parameters.xhtml#ipv6-parameters-2). This option is called "Per-Path Service Instruction Option". The "act" bits are 10 and the "chg" bit is 0. The suggested value is 144.

IANA is also requested to allocate a code point for Ethernet from the Assigned Internet Protocol Numbers registry (https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml). The suggested value is 143.

9. Acknowledgements

Thanks to Brian Carpenter, Adrian Farrel, Tom Herbert, John Leddy and Tony Li for their comments.

10. References

10.1. Normative References

[I-D.bonica-spring-srv6-plus]
Bonica, R., Hegde, S., Kamite, Y., Alston, A., Henriques, D., Jalil, L., Halpern, J., Linkova, J., and G. Chen, "Segment Routing Mapped To IPv6 (SRm6)", draft-bonica-spring-srv6-plus-06 (work in progress), October 2019.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

[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>.

[RFC4302] Kent, S., "IP Authentication Header", RFC 4302, DOI 10.17487/RFC4302, December 2005, <https://www.rfc-editor.org/info/rfc4302>.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, DOI 10.17487/RFC4303, December 2005, <https://www.rfc-editor.org/info/rfc4303>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, <https://www.rfc-editor.org/info/rfc4443>.

[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>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

10.2. Informative References

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, DOI 10.17487/RFC2784, March 2000, <https://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, <https://www.rfc-editor.org/info/rfc3031>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC4761] Kompella, K., Ed. and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, DOI 10.17487/RFC4761, January 2007, <https://www.rfc-editor.org/info/rfc4761>.

[RFC4762] Lasserre, M., Ed. and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, DOI 10.17487/RFC4762, January 2007, <https://www.rfc-editor.org/info/rfc4762>.

[RFC6624] Kompella, K., Kothari, B., and R. Cherukuri, "Layer 2 Virtual Private Networks Using BGP for Auto-Discovery and Signaling", RFC 6624, DOI 10.17487/RFC6624, May 2012, <https://www.rfc-editor.org/info/rfc6624>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

[RFC8077] Martini, L., Ed. and G. Heron, Ed., "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", STD 84, RFC 8077, DOI 10.17487/RFC8077, February 2017, <https://www.rfc-editor.org/info/rfc8077>.

Appendix A. Virtual Private Networks (VPN)

Virtual Private Network (VPN) technologies allow network providers to emulate private networks with shared infrastructure. For example, assume that red sites and blue sites connect to a provider network. The provider network facilitates communication among red sites and facilitates communication among blue sites. However, it prevents communication between red sites and blue sites.

The IETF has standardized many VPN technologies, including:

- Layer 2 VPN (L2VPN) [RFC6624].
- Layer 3 VPN (L3VPN) [RFC4364].
- Virtual Private LAN Service (VPLS) [RFC4761][RFC4762].
- Ethernet VPN (EVPN) [RFC7432].
- Pseudowires [RFC8077].

The above-mentioned technologies include the following components:

- Customer Edge (CE) devices.
- Provider Edge (PE) devices.
- Routing Instances.
- Service Instructions.
- Service Instruction Identifiers.

- Transport tunnels.

CE devices participate in closed communities called VPNs. CEs that participate in one VPN can communicate with one another but cannot communicate with CEs that participate in another VPN.

CE devices connect to provider networks through PE devices. Each PE maintains one Routing Instance for each VPN that it supports. A Routing Instance is a VPN specific Forwarding Information Base (FIB). In EVPN, Routing Instances are called Ethernet Virtual Instances (EVI).

Assume that one CE sends a packet through a provider network to another CE. The packet enters the provider network through an ingress PE and leaves the provider network through an egress PE. The packet may traverse one or more intermediate nodes on route from PE to PE.

When the ingress PE receives the packet, it:

- Identifies the Routing Instance that supports the originating CE's VPN.

- Searches that Routing Instance for the packet's destination.

If the search fails, the ingress PE discards the packet. If the search succeeds, it yields the following:

- A Service Instruction Identifier.

- The egress PE's IP address.

The ingress PE prepends the Service Instruction Identifier and a transport header to the packet, in that order. It then forwards the packet through a transport tunnel to the egress PE.

The egress PE removes the transport header, if it has not already been removed by an upstream device. It then examines and removes the Service Instruction Identifier. Finally, it executes a service instruction that is associated with the Service Instruction Identifier. The service instruction causes the egress PE to forward the packet to its destination (i.e., a directly connected CE).

In the above-mentioned VPN technologies, the ingress PE encodes Service Instruction Identifiers in Multiprotocol Label Switching (MPLS) [RFC3031] labels. Depending upon the transport tunnel type, the transport header can be:

- A MPLS label or label stack.
- An IPv4 [RFC0791] header.
- An IPv6 [RFC8200] header.
- A Generic Routing Encapsulation (GRE) [RFC2784] header encapsulated in IPv4 or IPv6.

Some PE devices cannot process MPLS headers. While these devices have several alternatives to MPLS-based transport tunnels, they require an alternative to MPLS-based encoding of Service Instruction Identifiers. The PPSI Option can be used to encode Service Instruction Identifiers . It is applicable when VPN payload is transported over IPv6.

Authors' Addresses

Ron Bonica
Juniper Networks
2251 Corporate Park Drive
Herndon, Virginia 20171
USA

Email: rbonica@juniper.net

Yuji Kamite
NTT Communications Corporation
3-4-1 Shibaura, Minato-ku
Tokyo 108-8118
Japan

Email: y.kamite@ntt.com

Luay Jalil
Verizon
Richardson, Texas
USA

Email: luay.jalil@one.verizon.com

Chris Lenart
Verizon
22001 Loudoun County Parkway
Ashburn, Virginia 20147
USA

Email: chris.lenart@verizon.com

Ning So
Reliance Jio
3010 Gaylord PKWY, Suite 150
Frisco, Texas 75034
USA

Email: Ning.So@ril.com

Fengman Xu
Reliance Jio
3010 Gaylord PKWY, Suite 150
Frisco, Texas 75034
USA

Email: Fengman.Xu@ril.com

Greg Presbury
Hughes Network Systems
11717 Exploration Lane
Germantown, Maryland 20876
USA

Email: greg.presbury@hughes.com

Gang Chen
Baidu
No.10 Xibeiwang East Road Haidian District
Beijing 100193
P.R. China

Email: phdgang@gmail.com

Yongqing Zhu
China Telecom
109 West Zhongshan Ave, Tianhe District
Guangzhou
P.R. China

Email: zhuyq.gd@chinatelecom.cn

Yifeng Zhou
ByteDance
Building 1, AVIC Plaza, 43 N 3rd Ring W Rd Haidian
District
Beijing 100000
P.R. China

Email: yifeng.zhou@bytedance.com