Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: February 26, 2022

Z. Li
L. Li
Huawei
H. Chen
Futurewei
C. Loibl
Next Layer Communications
G. Mishra
Verizon Inc.
Y. Fan
Casa Systems
Y. Zhu
China Telecom
L. Liu
Fujitsu
X. Liu
Volta Networks

August 25, 2021

BGP Flow Specification for SRv6
draft-li-idr-flowspec-srv6-07

Abstract

This document proposes extensions to BGP Flow Specification for SRv6 for filtering packets with a SRv6 SID that matches a sequence of conditions.

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 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/.

Copyright Notice

Table of Contents

1. Introduction

[RFC8955] describes in details about a new BGP NLRI to distribute a flow specification, which is an n-tuple comprising a sequence of matching criteria that can be applied to IP traffic. [RFC8956]

extends [RFC8955] to make it also usable and applicable to IPv6 data packets. [I-D.ietf-idr-flowspec-l2vpn] extends the flow-spec rules for layer 2 Ethernet packets. [I-D.hares-idr-flowspec-v2] specifies BGP Flow Specification Version 2.

Segment Routing (SR) for unicast traffic has been proposed to cope with the usecases in traffic engineering, fast re-reroute, service chain, etc. SR architecture can be implemented over an IPv6 data plane using a new type of IPv6 extension header called Segment Routing Header (SRH) [I-D.ietf-6man-segment-routing-header]. SRv6 Network Programming [RFC8986] defines the SRv6 network programming concept and its most basic functions. An SRv6 SID may have the form of LOC:FUNCT:ARG::.

LOC: Each operator is free to use the locator length it chooses. Most often the LOC part of the SID is routable and leads to the node which instantiates that SID.

FUNCT: The FUNCT part of the SID is an opaque identification of a local function bound to the SID. (e.g. End: Endpoint, End.X, End.T, End.DX2 etc.).

ARG: A function may require additional arguments that would be placed immediately after the FUNCT.

This document specifies one new BGP Flow Specification (FS) component type to support Segment Routing over IPv6 data plane (SRv6) filtering for BGP Flow Specification Version 2. The match field is destination address of IPv6 header, but it's a SRv6 SID from SRH rather than a traditional IPv6 address (refer to Figure 1). To support these features, a Flowspec version that is IPv6 capable (i.e., AFI = 2) MUST be used. These match capabilities of the features MAY be permitted to match when there is an accompanying SRH.

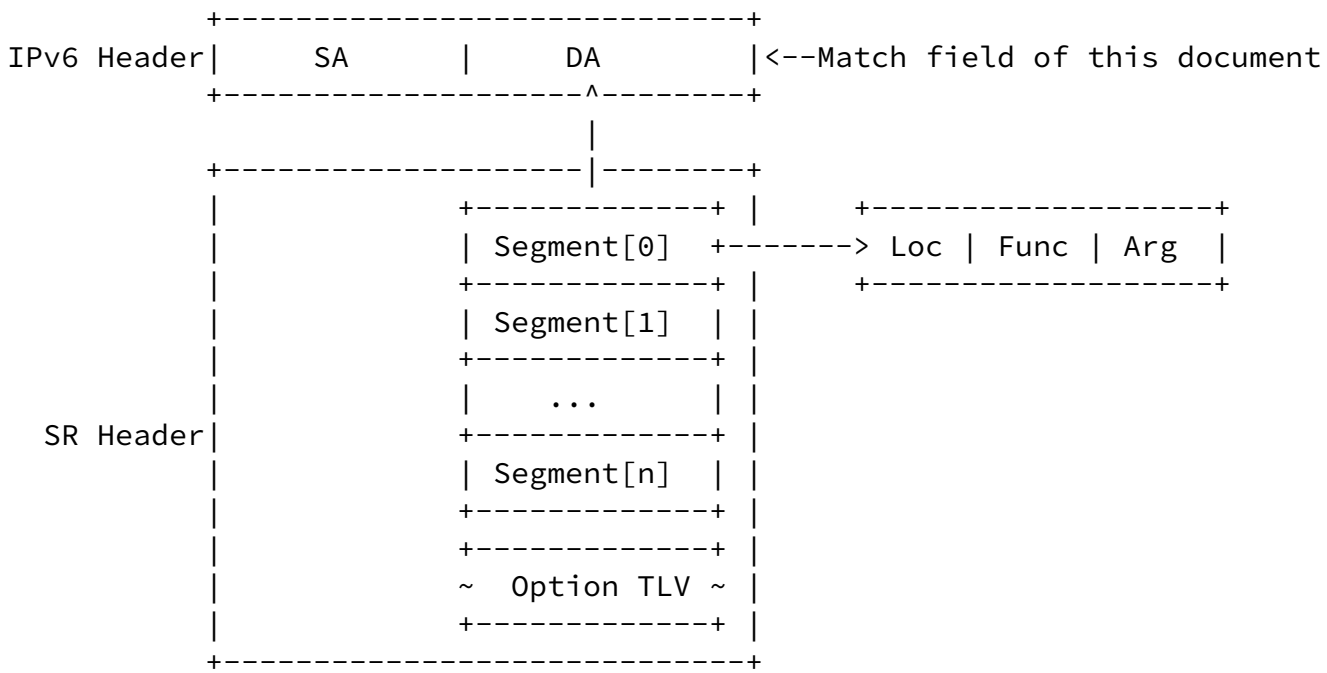


Figure 1: Match Field

2. Definitions and Acronyms

- FS: Flow Specification
- BGP-FS: Border Gateway Protocol (BGP) Flow Specification (FS)
- SR: Segment Routing

- SRH: SR Header.
- SRv6: IPv6 Segment Routing, SRv6 is a method of forwarding IPv6 packets on the network based on the concept of source routing.
- SID: Segment Identifier
- BSID: Binding SID

3. The Flow Specification Encoding for SRv6

The Flow Specification NLRI-type consists of several optional components, each of which begins with a type field (1 octet) followed by a variable length parameter. 13 component types are defined in [RFC8955] and [RFC8956] for IPv4 and IPv6. This document defines one component type for SRv6.

3.1. Type TBD1 - Some Parts of SID

[RFC8986] defines the format of SID is LOC:FUNCT:ARG::. In some scenarios, traffic packets can just match Locator, Function ID, Arguments or some combinations of these different fields. In order to match a part of SID, its prior parts need to be examined and matched first. For example, in order to match the Function ID (FUNCT), the Locator (LOC) needs to be examined and matched first. The new component type TBD1 defined below is for matching some parts of SID.

Encoding: <type, LOC-Len, FUNCT-Len, ARG-Len, [op, value]+>

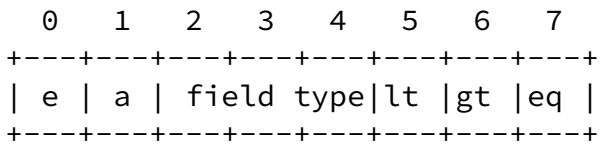
- type (1 octet): This indicates the new component type (TBD1, which is to be assigned by IANA).
- LOC-Len (1 octet): This indicates the length in bits of LOC in SID.
- FUNCT-Len (1 octet): This indicates the length in bits of FUNCT in SID.

o ARG-Len (1 octet): This indicates the length in bits of ARG in SID.

o [op, value]+: This contains a list of {operator, value} pairs that are used to match some parts of SID.

The total of three lengths (i.e., LOC length + FUNCT length + ARG length) MUST NOT be greater than 128. If it is greater than 128, an error occurs and Error Handling is applied according to [RFC7606] and [RFC4760].

The operator (op) byte is encoded as:



where the behavior of each operator bit has clear symmetry with that of [RFC8955]'s Numeric Operator field.

e - end-of-list bit. Set in the last {op, value} pair in the sequence.

a - AND bit. If unset, the previous term is logically ORed with the current one. If set, the operation is a logical AND. It should be

unset in the first operator byte of a sequence. The AND operator has higher priority than OR for the purposes of evaluating logical expressions.

field type:

000: SID's LOC

001: SID's FUNCT

010: SID's ARG

011: SID's LOC:FUNCT

100: SID's FUNCT:ARG

101: SID's LOC:FUNCT:ARG

For an unknown type, Error Handling is applied according to [RFC7606] and [RFC4760].

lt - less than comparison between data' and value'.

gt - greater than comparison between data' and value'.

eq - equality between data' and value'.

The data' and value' used in lt, gt and eq are indicated by the field type in a operator and the value field following the operator.

The value field depends on the field type and has the value of SID's some parts rounding up to bytes (refer to the table below).

Field Type	Value
SID's LOC	value of LOC bits
SID's FUNCT	value of FUNCT bits
SID's ARG	value of ARG bits
SID's LOC:FUNCT	value of LOC:FUNCT bits
SID's FUNCT:ARG	value of FUNCT:ARG bits
SID's LOC:FUNCT:ARG	value of LOC:FUNCT:ARG bits

3.2. Encoding Examples

3.2.1. Example 1

An example of a Flow Specification NLRI encoding for: all SRv6 packets to LOC 2001:db8:3::/48 and FUNCT {range [0100, 0300]}.

Some Parts of SID

```
                |
length          v               LOC==20010db80003  FUN>=100  FUN<=300
0x12          0f  30 10 40     01 2001 0db8 0003  4b 0100   bd 0300
                  ^  ^   ^
                  |  |   |
      Length of LOC FUN ARG
```

Decoded:

Value			
0x12	length	18 octets (if len<240, 1 octet)
TBD1(0x0f)	type	type TBD1(0x0f) - Some Parts of SID
0x30	LOC Length	= 48 (bits)
0x10	FUNCT Length	= 16 (bits)
0x40	ARG Length	= 64 (bits)
0x01	op	LOC ==
0x2001	value	LOC's value = 2001:db8:3
0x0db8		
0x0003		
0x4b	op	"AND", FUNCT >=
0x0100	value	FUNCT's value = 0100
0xbd	op	end-of-list, "AND", FUNCT <=
0x0300	value	FUNCT's value = 0300

4. Security Considerations

No new security issues are introduced to the BGP protocol by this specification over the security considerations in [RFC8955] and [RFC8956].

5. IANA Considerations

Under "Flow Spec Component Types" registry, IANA is requested to assign the following values:

Value	IPv4 Name	IPv6 Name	Reference
TBD1	Unassigned	Some Parts of SID	This Document

6. Acknowledgments

The authors would like to thank Joel Halpern, Jeffrey Haas, Ketan Talaulikar, Aijun Wang, Dhruv Dhody, Shunwan Zhuang and Rainsword Wang for their valuable suggestions and comments on this draft.

7. References

7.1. Normative References

[I-D.hares-idr-flowspec-v2]
Hares, S. and D. Eastlake, "BGP Flow Specification Version 2", draft-hares-idr-flowspec-v2-02 (work in progress), July 2021.

[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>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[RFC7153] Rosen, E. and Y. Rekhter, "IANA Registries for BGP Extended Communities", RFC 7153, DOI 10.17487/RFC7153, March 2014, <https://www.rfc-editor.org/info/rfc7153>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

[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>.

[RFC8955] Loibl, C., Hares, S., Raszuk, R., McPherson, D., and M. Bacher, "Dissemination of Flow Specification Rules", RFC 8955, DOI 10.17487/RFC8955, December 2020, <https://www.rfc-editor.org/info/rfc8955>.

[RFC8956] Loibl, C., Ed., Raszuk, R., Ed., and S. Hares, Ed., "Dissemination of Flow Specification Rules for IPv6", RFC 8956, DOI 10.17487/RFC8956, December 2020, <https://www.rfc-editor.org/info/rfc8956>.

7.2. Informative 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-idr-flowspec-l2vpn]
Hao, W., Eastlake, D. E., Litkowski, S., and S. Zhuang, "BGP Dissemination of L2 Flow Specification Rules", draft-ietf-idr-flowspec-l2vpn-17 (work in progress), May 2021.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

Authors' Addresses

Zhenbin Li
Huawei
156 Beiqing Road
Beijing, 100095
P.R. China

Email: lizhenbin@huawei.com

Lei Li
Huawei
156 Beiqing Road
Beijing 100095
P.R. China

Email: lily.lilei@huawei.com

Huaimo Chen
Futurewei
Boston, MA
USA

Email: Huaimo.chen@futurewei.com

Christoph Loibl
Next Layer Communications
Mariahilfer Guertel 37/7
Vienna 1150
AT

Email: cl@tix.at

Gyan S. Mishra
Verizon Inc.
13101 Columbia Pike
Silver Spring MD 20904
USA

Phone: 301 502-1347
Email: gyan.s.mishra@verizon.com

Yanhe Fan
Casa Systems
USA

Email: yfan@casa-systems.com

Yongqing Zhu
China Telecom
109, West Zhongshan Road, Tianhe District
Guangzhou 510000
China

Email: zhuyq8@chinatelecom.cn

Lei Liu
Fujitsu
USA

Email: liulei.kddi@gmail.com

Xufeng Liu
Volta Networks
McLean, VA
USA

Email: xufeng.liu.ietf@gmail.com