Workgroup: Network Working Group
Internet-Draft:
draft-ietf-idr-ts-flowspec-srv6-policy-02
Published: 6 March 2023
Intended Status: Informational
Expires: 7 September 2023
Authors: W. Jiang, China Mobile
Y. Liu, China Mobile
S. Zhuang, Huawei Technologies
G. Mishra, Verizon Communications Inc.
S. Chen, Huawei Technologies

Traffic Steering using BGP FlowSpec with SR Policy

Abstract

BGP Flow Specification (FlowSpec) [RFC8955] [RFC8956] has been proposed to distribute BGP FlowSpec NLRI to FlowSpec clients to mitigate (distributed) denial-of-service attacks, and to provide traffic filtering in the context of a BGP/MPLS VPN service. Recently, traffic steering applications in the context of SR-MPLS and SRv6 using FlowSpec also attract attention. This document introduces the usage of BGP FlowSpec to steer packets into an SR Policy.

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].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

SR-MPLS [RFC8660] forwards data packets using the source routing model. The core idea of SR-MPLS is to divide a packet forwarding path into different segments, allocate segment identifiers (SIDs) to the segments, and encapsulate segment information into packets at the ingress of the path to guide packet forwarding.

Segment Routing IPv6 (SRv6) is a protocol designed to forward IPv6 data packets on a network using the source routing model. SRv6 enables the ingress network device to add a segment routing header (SRH) [RFC8754] to an IPv6 packet and push an explicit IPv6 address stack into the SRH. After receiving the packet, each transit node updates the IPv6 destination IP address in the packet and segment list to implement hop-by-hop forwarding.

SR Policy (includes SR-MPLS and SRv6 Policy) [RFC9256] is a tunneling technology developed based on SR-MPLS and SRv6. An SR Policy is a set of candidate paths consisting of one or more segment lists, that is, segment ID (SID) lists. Each SID list identifies an end-to-end path from the source node to the destination node, instructing a network device to forward traffic through the path rather than the shortest path computed using an IGP. The header of a packet steered into an SR Policy is augmented with an ordered list of segments associated with that SR Policy, so that other devices on the network can execute the instructions encapsulated into the list.

The headend of an SR Policy may learn multiple candidate paths for an SR Policy. Candidate paths may be learned via a number of different mechanisms, e.g., CLI, NetConf, PCEP[I-D.ietf-pce-segment-routing-policy-cp], or BGP[I-D.ietf-idr-segment-routing-te-policy].

[RFC8955] [RFC8956] defines the flow specification (FlowSpec) that allows to convey flow specifications and traffic Action/Rules associated (rate- limiting, redirect, remark ...). BGP Flow specifications are encoded within the MP_REACH_NLRI and MP_UNREACH_NLRI attributes[RFC4760]. Rules (Actions associated) are encoded in Extended Community attribute[RFC4360]. The BGP Flow Specification function allows BGP Flow Specification routes that carry traffic policies to be transmitted to BGP Flow Specification peers to steer traffic.

This document proposes BGP flow specification usage that are used to steer data flow into an SR Policy as well as to indicate Tailend function for SRv6 scenario. This work is helpful for promoting the deployment of SR-MPLS and SRv6 networks.

2. Definitions and Acronyms

*FlowSpec: Flow Specification

*SR: Segment Routing

*SR-MPLS: SR over the MPLS data plane

*SRv6: SR over the IPv6 data plane

*SID: Segment Identifier

*SRH: Segment Routing Header

*TE: Traffic Engineering

*USD: Ultimate Segment Decapsulation

3. Operations

An SR Policy [RFC9256] is identified through the tuple <headend, color, endpoint>. In the context of a specific headend, one may identify an SR Policy by the <color, endpoint> tuple. The headend is the node where the SR Policy is instantiated/implemented. The headend is specified as an IPv4 or IPv6 address and is expected to be unique in the domain. The endpoint indicates the destination of the SR Policy. The endpoint is specified as an IPv4 or IPv6 address and is expected to be unique in the domain. The color is a 32-bit unsigned numerical value that associates with the SR policy, and it defines an application-level network Service Level Agreement (SLA) policy or intent.

Assume one or multiple SR Policies are already setup/instantiated in the SR HeadEnd device. In order to steer traffic into a specific SR Policy at the Headend, one can use the SR Color Extended community [RFC9012] and endpoint to map to a satisfying SR Policy, and steer the traffic into this specific SR Policy.

[I-D.ietf-idr-flowspec-redirect-ip] defines the redirect to IPv4 and IPv6 Next-hop action. The IPv4 next-hop address in the Flow-spec Redirect to IPv4 Extended Community can be used to specify the endpoint of the SR Policy, and the IPv6 next-hop address in the Flow-spec Redirect to IPv6 Extended Community[RFC5701] can be used to specify the endpoint of the SRv6 Policy. When the packets reach to the TailEnd device, some specific function information identifiers can be used to decide how to further process the flows in SRv6 scenario. Several endpoint functions are already defined, e.g., End.DT6: Endpoint with decapsulation and IPv6 table lookup, and End.DX6: Endpoint with decapsulation and IPv6 cross-connect. The BGP Prefix-SID defined in [RFC8669] is utilized to enable SRv6 VPN services [RFC9252]. SRv6 Services TLVs within the BGP Prefix-SID Attribute can be used to indicate the endpoint functions.

For SR-MPLS scenario, this document proposes to carry the Color Extended Community and the Flow-spec Redirect to IPv4 Extended Community in the context of a Flowspec NLRI [RFC8955] [RFC8956] to an SR-MPLS Headend to steer traffic into one SR-MPLS Policy.

For SRv6 scenario, this document proposes to carry the Color Extended Community, the Flow-spec Redirect to IPv6 Extended Community and BGP Prefix-SID Attribute in the context of a Flowspec NLRI [RFC8955] [RFC8956] to an SRv6 Headend to steer traffic into one SRv6 Policy, as well as to indicate specific Tailend functions.

For the case that a flowspec route carries multiple Color Extend Communities, the Color Extended community with the highest numerical value will be given higher preference per the description in Section 8.4.1 of [RFC9256].

The method proposed in this document supports load balancing to the tailend device. To achieve that, the headend device CAN set up multiple paths in one SR Policy, and use a Flowspec route to indicate the specific SR Policy.

4. SR-MPLS Application Examples

In following scenario, BGP FlowSpec Controller signals the filter rules, the Flow-spec Redirect to IPv4 action, and the policy color to the SR-MPLS HeadEnd device.

```
   +------------+
   |  BGP FS    |
   | Controller |
   +------------+
      | FlowSpec route to HeadEnd:
      |   NLRI: Filter Rules
      |   Redirect to IPv4 Nexthop: TailEnd's Address
      |   Policy Color: C0
      |
      |          .------.
      |         (        )
      V     .--(          )--.
+-------+  (                  )  +-------+
|       |_(  SR-MPLS Network  )_|       |
|HeadEnd| ( ================> ) |TailEnd|
+-------+  (SR List<S1,S2,S3>)  +-------+
            '--(          )--'
                (        )
                 '-----'
```

Figure 1: Steering the Traffic Flow into SR-MPLS Policy

When the SR-MPLS HeadEnd device (as a FlowSpec client) receives such instructions from BGP FS Controller, it will steer the traffic flows matching the criteria in the FlowSpec route into the SR-MPLS Policy matching the tuple (Endpoint: TailEnd's Address, Color: C0). And the packets of such traffic flows will be encapsulated with MPLS stack using the SR List <S1, S2, S3> in the HeadEnd device, then send the packets to the TailEnd device along the path indicated by the SR list.

5. SRv6 Application Examples

In following scenario, BGP FlowSpec Controller signals the filter rules, the redirect to IPv6 Nexthop action, the policy color and the function information (SRv6 SID: Service_id_x) to the HeadEnd device.

```
   +------------+
   |  BGP FS    |
   | Controller |
   +------------+
      | FlowSpec route to HeadEnd:
      |   NLRI: Filter Rules
      |   Redirect to IPv6 Nexthop: TailEnd's Address
      |   Policy Color: C1
      |   PrefixSID: Service_id_x
      |          .-----.
      |         (       )
      V     .--(         )--.
+-------+  (                 )  +-------+
|       |_( SRv6 Core Network )_|       |
|HeadEnd| ( ================> ) |TailEnd|
+-------+  (SR List<S1,S2,S3>)  +-------+
            '--(         )--'   Service SID: Service_id_x
                (       )       (e.g.: End.DT4 or End.DT6 or others)
                 '-----'
```

Figure 2: Steering the Traffic Flow into SRv6 Policy (Option 1)

When the HeadEnd device (as a FlowSpec client) receives such instructions from BGP FS Controller, it will steer the traffic flows matching the criteria in the FlowSpec route into the SRv6 Policy matching the tuple (Endpoint: TailEnd's Address, Color: C1). And the packets of such traffic flows will be encapsulated with SRH(Segment Routing Header) using the SR List <S1, S2, S3, Service_id_x>. When the packets reach to the TailEnd device, they will be further processed per the function denoted by the Service_id_x.

When the HeadEnd device determines (with the help of SRv6 SID Structure) that the Service SID belongs to the same SRv6 Locator as the last SRv6 SID of the TailEnd device in the SRv6 Policy segment list, it MAY exclude that last SRv6 SID when steering the service flow. For example, the effective segment list of the SRv6 Policy associated with SID list <S1, S2, S3> would be replaced as <S1, S2, Service_id_x>.

If the last SRv6 SID (For example, S3 we use here) of the TailEnd device in the SRv6 Policy segment list is USD-flavored, an SRv6 Service SID (e.g., End.DT4 or End.DT6) is not required when BGP

FlowSpec Controller sends the FlowSpec route to the HeadEnd device (as a FlowSpec client).

```
   +------------+
   |  BGP FS    |
   | Controller |
   +------------+
      | FlowSpec route to HeadEnd:
      |   NLRI: Filter Rules
      |   Redirect to IPv6 Nexthop: TailEnd's Address
      |   Policy Color: C2
      |          .-----.
      |         (       )
      V     .--(         )--.
+-------+  (                 )  +-------+
|       |_( SRv6 Core Network )_|       |
|HeadEnd| ( ================> ) |TailEnd|
+-------+  (SR List<S1,S2,S3>)  +-------+
            '--(         )--'
                (       )
                 '-----'
Note: S3 MUST be a USD-flavored SRv6 SID of the TailEnd
```

Figure 3: Steering the Traffic Flow into SRv6 Policy (Option 2)

When the HeadEnd device (as a FlowSpec client) receives such instructions from BGP FS Controller, it will steer the traffic flows matching the criteria in the Flowspec route into the SRv6 Policy matching the tuple (Endpoint: TailEnd's Address, Color: C2). And the packets of such traffic flows will be encapsulated with SRH(Segment Routing Header) using the SR List <S1, S2, S3>. When the packets reach to the TailEnd device, they will be further processed per the function denoted by the USD-flavored SRv6 SID S3.

At this point, the work discusses the matching of global routing table prefixes.

For the cases of intra-AS and inter-AS traffic steering using this method, the usages of Flowspec Color Extended Community with BGP prefix SID are the same for both scenarios. The difference lies between the local SRv6 policy configurations. For the inter-domain case, the operator can configure an inter-domain SRv6 policy/path at the Headend device. For the intra-domain case, the operator can configure an intra-domain SRv6 policy/path at the Headend device.

6. Running Code

6.1. Interop-test Status

The Traffic Steering using BGP FlowSpec with SR-MPLS / SRv6 Policy mechanism has been implemented on the following hardware devices, Network Operating System software and SDN controllers. They had also successfully participated in the series of joint interoperability testing events hosted by China Mobile from July 2021 to October 2021. The following hardware devices and Network Operating System software had successfully passed the interoperability testing (in alphabetical order).

Routers:

Vendors	Device Model	Version
Huawei	NE40-X8A	NE40E V800R021C00SPC091T
New H3C	CR16010H-FA	Version 7.1.075, ESS 8305
Ruijie	RG-N8010-R	N8000-R_RGOS 12.8(1)B08T1
ZTE	M6000-8S Plus	V5.00.10(5.60.5)

Controllers:

Vendors	Device Model	Version
China Unitechs	I-T-E SC	V1.3.6P3
Huawei	NCE-IP	V100R021C00
Ruijie	RG-ONC-AIO-H	RG-ION-WAN-CLOUD_2.00T1
ZTE	ZENIC ONE	R22V16.21.20

6.2. Deployment Status

Currently, this feature has been deployed on the IP backbone network of China Mobile.

7. IANA Considerations

No IANA actions are required for this document.

8. Security Considerations

This document does not change the security properties of SRv6 and BGP.

9. Contributors

The following people made significant contributions to this document:

Yunan Gu
Huawei Technologies
Email: guyunan@huawei.com

Haibo Wang
Huawei Technologies
Email: rainsword.wang@huawei.com

Jie Dong
Huawei Technologies
Email: jie.dong@huawei.com

Xue Yang
China Mobile
Email: yangxuewl@chinamobile.com

10. Acknowledgements

The authors would like to acknowledge the review and inputs from Jeffrey Haas, Susan Hares, Weiqiang Cheng, Kaliraj Vairavakkalai, Robin Li, Acee Lindem, Gunter Van De Velde, John Scudder, Rainbow Wu, Linda Dunbar, Gang Yan, Feng Yang, Wim Henderickx, Robert Raszuk, Ketan Talaulikar, Changwang Lin, Aijun Wang, Hao Li, Huaimo Chen, Sheng Fang, Yuanxiang Qiu, Ran Chen, Cheng Li, Zheng Zhang, Xuewei Wang, Yanrong Liang, Xuhui Cai, Haojie Wang, Lili Wang and Nan Geng.

Special thanks to Nat Kao, who suggested adding SR-MPLS use cases to this document.

11. References

11.1. Normative References

[I-D.ietf-idr-flowspec-redirect-ip]
Uttaro, J., Haas, J., Texier, M., akarch@cisco.com, Ray, S., Simpson, A., and W. Henderickx, "BGP Flow-Spec Redirect to IP Action", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-redirect-ip-02, 5 February 2015,

<https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-redirect-ip-02>.

[I-D.ietf-idr-segment-routing-te-policy]
Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., Jain, D., and S. Lin, "Advertising Segment Routing Policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-segment-routing-te-policy-20, 27 July 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-segment-routing-te-policy-20>.

[I-D.ietf-pce-segment-routing-policy-cp]
Koldychev, M., Sivabalan, S., Barth, C., Peng, S., and H. Bidgoli, "PCEP extension to support Segment Routing Policy Candidate Paths", Work in Progress, Internet-Draft, draft-ietf-pce-segment-routing-policy-cp-08, 24 October 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-pce-segment-routing-policy-cp-08>.

[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>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <https://www.rfc-editor.org/info/rfc4360>.

[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>.

[RFC5701] Rekhter, Y., "IPv6 Address Specific BGP Extended Community Attribute", RFC 5701, DOI 10.17487/RFC5701, November 2009, <https://www.rfc-editor.org/info/rfc5701>.

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

[RFC8669] Previdi, S., Filsfils, C., Lindem, A., Ed., Sreekantiah, A., and H. Gredler, "Segment Routing Prefix Segment

Identifier Extensions for BGP", RFC 8669, DOI 10.17487/RFC8669, December 2019, <https://www.rfc-editor.org/info/rfc8669>.

[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>.

[RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <https://www.rfc-editor.org/info/rfc9012>.

[RFC9252] Dawra, G., Ed., Talaulikar, K., Ed., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)", RFC 9252, DOI 10.17487/RFC9252, July 2022, <https://www.rfc-editor.org/info/rfc9252>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

11.2. Informative References

[I-D.ietf-idr-flowspec-path-redirect] Van de Velde, G., Patel, K., and Z. Li, "Flowspec Indirection-id Redirect", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-path-redirect-12, 24 November 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-path-redirect-12>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

Authors' Addresses

Wenying Jiang
China Mobile
No.32 XuanWuMen West Street
Beijing
100053
China

Email: jiangwenying@chinamobile.com

Yisong Liu
China Mobile
No.32 XuanWuMen West Street
Beijing
100053
China

Email: liuyisong@chinamobile.com

Shunwan Zhuang
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing
100095
China

Email: zhuangshunwan@huawei.com

Gyan Mishra
Verizon Communications Inc.
13101 Columbia Pike
Silver Spring, MD 20904,
United States of America

Email: gyan.s.mishra@verizon.com

Shuanglong Chen
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing
100095
China

Email: chenshuanglong@huawei.com