

IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: 17 September 2025

H. Zhang
Z. Huang
C. Huang
China Telecom
16 March 2025

BGP FlowSpec Extension for Traffic Scheduling
draft-zhang-idr-bgp-flowspec-extension-01

Abstract

Traditional QoS technology can not achieve fine adjustment in the traffic scheduling, and has a large amount of configuration and poor maintainability. BGP FlowSpec technology provides a wealth of filtering conditions and processing actions, using BGP network layer reachable information (NLRI) to transmit traffic filtering information, which can achieve a more fine-grained control of the traffic and improve maintainability.

This document defines the extension of BGP FlowSpec, adding two new traffic filtering actions for extended community: minimum rate guarantee and congestion management, to enable better management and scheduling of traffic, and further improving the scalability and applicability of BGP FlowSpec.

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 17 September 2025.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

| | |
|--|---|
| 1. Introduction | 2 |
| 1.1. Terminology | 2 |
| 2. Extended Community Flow Specification Actions | 3 |
| 3. Use Cases | 4 |
| 4. IANA Considerations | 5 |
| 5. References | 5 |
| Authors' Addresses | 6 |

1. Introduction

[RFC8955] and [RFC8956] have defined commonly used traffic filtering conditions and actions for IPv4 and IPv6. The traffic filtering conditions are used as NLRI and the traffic filtering actions are used as Extended Community in BGP FlowSpec routing.

[I-D.ietf-idr-flowspec-v2] defines BGP Flow Specification Version 2. More refined traffic management can be achieved by traffic scheduling that matches the rules through the traffic filtering conditions and the corresponding actions carried in BGP FlowSpec routes.

This document defines two new extended community for traffic filtering actions of BGP flowspec, which respectively represent the minimum rate guarantee and congestion management for the flow, in order to better realize traffic scheduling through BGP FlowSpec routes.

1.1. Terminology

NLRI: Network Layer Reachability Information

DSCP: Diffserv Code Point

2. Extended Community Flow Specification Actions

This document defines two new extended community: minimum rate guarantee and congestion management.

| Type | Value | Action | Encoding |
|------|------------------------------|--------|--------------------------|
| TBD1 | CBR (Minimum rate guarantee) | | 2-octet AS, 4-octet Rate |
| TBD2 | CM (Congestion Management) | | 1-octet Queue value |

The specific encoding for the first new extended community is shown as follows:

```

 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           TBD1           |           ASN           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                           Rate                           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where,

Type: 2 bytes, representing the minimum rate guarantee attribute, to be allocated by IANA.

ASN: 2 bytes, representing the AS number.

Rate: 4 bytes, representing the minimum guaranteed rate, expressed in IEEE floating point format, units being bytes per second. The router sets the traffic flow as fast forwarding based on this attribute, ensuring that the CBR is the Rate value.

The specific encoding for the second new extended community defined in this document is shown as follows:

```

 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           TBD2           |           Queue           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where,

Type: 2 bytes, representing congestion management attribute, to be allocated by IANA.

Queue: 2 byte, indicating the type used for queue scheduling. The lower three bits have values from 0 to 7, representing queues BE, AF1, AF2, AF3, AF4, EF, CS6, and CS7, corresponding to the commonly used 8 queues currently, and the other bits are reserved. After the router is configured with congestion management mechanism (which may be queue technologies such as FIFO, PQ, CQ, RTPQ, WFQ, CBQ, CBR, etc., not discussed in this document), traffic of different queue types will be forwarded according to the corresponding queue scheduling technology.

3. Use Cases

This section describes how to use the defined new extended community in real scenarios.

Example 1:

The traffic with a destination address of 2.2.2.0/24 and DSCP of 3 requires a guaranteed minimum rate of 30 Mb/s. The BGP FlowSpec controller advertises a BGP route to the Ingress Router, carrying the following information.

Flow Specification NLRI encoding:

| Length | Destination | DSCP |
|--------|----------------|----------|
| 0x08 | 01 18 02 02 02 | 0B 81 0C |

Extended Community:

| Type | ASN | Rate |
|------|-------|-------------|
| TBD1 | 00 00 | 01 c9 c3 80 |

Parsing the NLRI encoding of the above route as the traffic filtering condition: the destination address is 2.2.2.0/24 and the DSCP is 3. Traffic that matches this condition will perform the action carried in the Extended Community: minimum traffic rate guarantee of 30M, so as to achieve traffic scheduling.

Example 2:

The traffic with a destination address of 192.0.2.0/24 and DSCP of 3 is set to AF2 queue. The router forwards the message according to its configured queue scheduling method. The BGP FlowSpec controller advertises a BGP route to the Ingress Router, carrying the following information.

Flow Specification NLRI encoding:

| Length | Destination | DSCP |
|--------|----------------|----------|
| 0x08 | 01 18 c0 00 02 | 0B 81 0C |

Extended Community:

| Type | Queue |
|------|-------|
| TBD1 | 00 02 |

Parsing the NLRI encoding of the above route as the traffic filtering condition: the destination address is 192.2.2.0/24 and the DSCP is 3. Traffic that matches this condition will perform the action carried in the Extended Community: traffic queue is AF2, so as to achieve traffic scheduling.

4. IANA Considerations

For the purpose of this work, the following two Extended Communities require IANA assignments.

TBD1 for CBR (Minimum rate guarantee, bytes).

TBD2 for CM (Congestion Management).

5. References

[I-D.ietf-idr-flowspec-v2]

Hares, S., Eastlake, D. E., Yadlapalli, C., and S. Maduschke, "BGP Flow Specification Version 2", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-v2-04, 28 April 2024, <<https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-v2-04>>.

[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>>.
- [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>>.

Authors' Addresses

Huiyue Zhang
China Telecom
Email: zhanghy30@chinatelecom.cn

Zhuojun Huang
China Telecom
Email: huangzhuoj@chinatelecom.cn

Cancan Huang
China Telecom
Email: huangcanc@chinatelecom.cn