IDR
Internet-Draft
Intended status: Informational
Expires: 10 November 2026

Y. Cui
Tsinghua University
Y. Gao
Zhongguancun Laboratory
S. Hares
Hickory Hill Consulting

9 May 2026

Packet Content Filter for BGP FlowSpec
draft-cui-idr-content-filter-flowspec-04

Abstract

The BGP Flow Specification enables the distribution of traffic filter policies (traffic filters and actions) via BGP, facilitating DDoS traffic filtering. However, the traffic filter in FSv1 and FSv2 predominantly focuses on IP header fields, which may not adequately address volumetric DDoS attack traffic characterized by fixed patterns within the packet content. This document introduces a new flow specification filter type designed for packet content filtering. The match field includes ptype, otype, offset, content-length, content, and mask encoded in the Flowspec NLRI. This new filter aims to leverage network devices such as routers and switches to support controlled traffic handling, traffic optimization, and mitigation of simple volumetric DDoS attacks, reducing the overall processing cost of carrier networks.

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 10 November 2026.

Table of Contents

1. Introduction

BGP flow specification describes the distribution of traffic filter policies through BGP, allowing for efficient traffic management and DDoS attack mitigation. Existing versions, FSv1 and FSv2, primarily offer n-tuple matching conditions for policy enforcement, enabling actions such as packet dropping, re-directing, or other actions. These filter rules can be propagated to all BGP peers simultaneously without necessitating router configuration changes. Despite their utility, the reliance of existing filters on IP header fields may be insufficient for some operational scenarios where packets can be identified by fixed content patterns. Such scenarios may include

DDoS mitigation, traffic filtering, and traffic optimization. Some attacks or traffic classes may contain fixed patterns in the packet payload that can be matched at known offsets.

This document defines a new FlowSpec filter type that supports packet content filtering by using ptype, otype, offset, content-length, content, and mask fields within the FlowSpec NLRI. This filter is intended for controlled operational use cases such as traffic filtering, traffic optimization, and DDoS mitigation.

1.1. Terminology

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.

2. Definitions and Acronyms

* DDoS: Distributed Denial of Service.

* NLRI: Network Layer Reachability Information.

* FSv1: Flow Specification Version 1, defined in [RFC8955] and [RFC8956].

* FSv2: Flow Specification Version 2, defined in [I-D.ietf-idr-flowspec-v2].

3. The Packet Content Filter for FSv1

This document specifies a new flow specification filter type that is encoded in the BGP FS NLRI, following the FSv1 definition format. The packet content filter is defined as follows:

Type TBD Packet-Content

Encoding:< type (1 octet), value>

The value field is encoded using ptype, otype, offset, content-length, content and mask.

Encoding: < ptype (4 bits), otype (4 bits), offset (2 octets), content-length (1 octet), content (Variable), mask (Variable)>

3.1. Ptype Field

The ptype is defined as a 4-bit unsigned integer that defines the packet type via AFI, because some filters are added to hardware that are IPv4 or IPv6 specific.

Value	Description of Ptype
1	IPv4
2	IPv6

Figure 1: Ptype field

3.2. Otype and Offset Fields

The otype and offset fields define the starting position of the packet content used for matching.

To avoid the effect of variable header length on the offset, we use the hierarchical way like [I-D.khare-idr-bgp-flowspec-payload-match]. The Otype is defined as a 4-bit unsigned integer. The detail are as follows:

Value	Description of Otype
0	IP Header
1	IP Payload
2	UDP Payload
3	TCP Payload

Figure 2: Otype field

Otype 0 is defined as the start of the IP header. Otype 1 is defined as the start of the data portion of the IP header after the IP options. Otype 2 is defined as the start of the UDP payload. Otype 3 is defined as the start of the TCP payload. Otype 2 MUST only match packets whose upper-layer protocol is UDP (17). Otype 3 MUST only match packets whose upper-layer protocol is TCP (6). For other IP protocols, otype 2 and otype 3 MUST NOT match; otype 1 MAY be used.

The offset is defined as a 2-octet unsigned integer that specifies the count of octets to be bypassed from the otype’s starting position to match the packet content. It is worth noting that packet

fragmentation will cause the offset value to change, so it is not enough to filter the fragmented packets through the packet content filter. One possible way is to filter the first packet through the payload filter, and then use its header information along with the fragment filter to filter the subsequent packets.

Example:

* By setting otype 0 and an offset of 0, the match is configured to start precisely at the beginning of the IP header.

* By setting otype 1 and an offset of 2, the match will start two octets past the initial data portion of the IP header, skipping over any IP options. This configuration, for example, could be used to specifically target the IP payload starting after 2 octets.

* By setting otype 2 and an offset of 10, the match will start ten octets into the UDP payload of the packet.

* By setting otype 3 and an offset of 10, the match will start ten octets into the TCP payload of the packet.

3.3. Content-length, Content and Mask Fields

The content-length is a one-octet unsigned integer field that specifies the length, in octets, of each of the Content field and the Mask field. The Content field and the Mask field have the same length as specified by the content-length.

The Content field carries the octet sequence to be matched. Based on information provided by equipment vendors and operators, 8 octets is usually sufficient for identifying many fixed packet-content patterns used in operational filtering scenarios.

The Mask field is an octet string used as a bit mask for the Content field and the corresponding packet data. Each bit set to 1 indicates that the corresponding bit is significant for matching. Each bit set to 0 indicates that the corresponding bit is ignored.

A packet matches the Packet Content filter if the following comparison is true for the packet data at the specified offset:
(packet_content & mask) == (content & mask)

3.4. Example of Encoding

An example of a FlowSpec NLRI encoding is provided for the following rule: "match all packets destined to 192.0.2.0/24 that have the fixed content 0x5858 at offset 0 in the TCP payload".

length	destination	packet content
0x0f	01 18 c0 00 02	TBD 40 12 00 00 02 58 58 ff ff

Table 1

Description of each field of the FlowSpec NLRI.

Value	Description	
0x0f	length	15 octets (if len<240, 1 octet)
0x01	type	Type 1 - Destination Prefix
0x18	length	24 bits
0xc0	prefix	192
0x00	prefix	0
0x02	prefix	2
TBD	type	Type TBD - Packet Content
0x40	length	64 bits
0x12	ptype, otype	IPv4, TCP payload
0x0000	offset	0 octets
0x02	content-length	2 octets
0x5858	content	0x5858
0xffff	mask	0xffff

Table 2

4. The Packet Content Filter for FSv2

4.1. Filter Encoding

To adapt to the updates of FlowSpec, this document also defines the Packet Content Filter for FSv2. The format follows the NLRI format for Extended IP Filters defined in [I-D.hares-idr-fsv2-more-ip-filters], as shown in Figure 3:

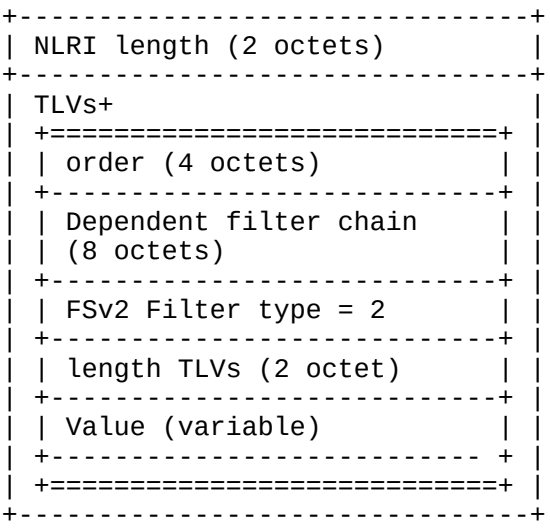

Figure 3: NLRI Format for Extended IP Filters

The format of the dependent filter chain is shown in Figure 4:

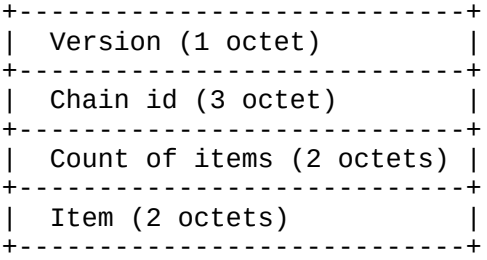

Figure 4: Format of the Dependent Filter Chain

The format of the Component TLV is shown in Figure 5:

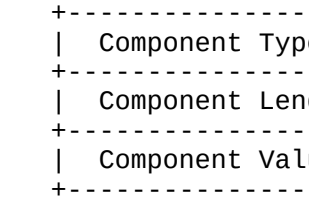

Figure 5: format for Component-TLV

The definition of the Packet Content Filter in the Component-TLV format is as follows:

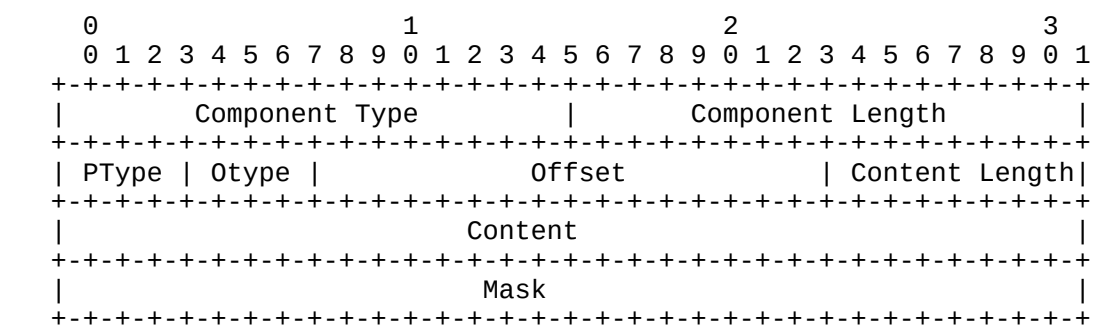

Figure 6: Definition of the Packet Content Filter

where the fields have the same definitions as in the FSv1 encoding.

Encoding: < ptype (4 bits), otype (4 bits), offset (2 octets), content-length (1 octet), content (Variable), mask (Variable)>

4.2. Filter Ordering Rule

Compared to FSv1, FSv2 adds filter ordering function. According to the definition of ordering rules in FSv2, the transmission of Component-TLVs within a flow specification rule MUST be sent ascending order by Component-TLV type. If the Component-TLV types are the same, then the value fields are compared using mechanisms defined in [RFC8955] and [RFC8956] and MUST be in ascending order.

However, due to multiple fields in the value of the packet content filter, the mechanisms defined in [RFC8955] and [RFC8956] do not apply. To give the default ordering rules of packet content filters, this document gives the definition as follows:

1. Filters with a larger content-length are ordered first.

2. If they have the same content-length, compare otype and the larger type is ordered first.

3. If they have the same content-length and otype, compare offset and the larger value is ordered first.

4. If they have the same content-length, otype, and offset, compare the content as an unsigned octet string in lexicographic order, starting from the first octet. If the common prefix is not equal, the string with the lower octet value at the first differing position has higher precedence.

When multiple Packet Content Filter components exist across multiple NLRIs with the same user order, their relative order is determined according to the ordering rules above.

4.3. Use Cases

Here is a use case for ordering rules with multiple NLRI and multiple components. There are five components, with the same destination IP and user order, each of which contains a packet content filter with different values:

User-Order 10

FSv2 NLRI with Extended IP Filters

Component 1: Destination IP + Packet content filter (otype 0, offset 50, content-length 2 , content 0x1111) + Rate Limit

Component 2: Destination IP + Packet content filter (otype 0, offset 50, content-length 3 , content 0x111122) + Discard

Component 3: Destination IP + Packet content filter (otype 2, offset 70, content-length 2 , content 0x1111) + Rate Limit

Component 4: Destination IP + Packet content filter (otype 2, offset 70, content-length 3 , content 0x111122) + Discard

Component 5: Destination IP + Packet content filter (otype 2, offset 70, content-length 3 , content 0x111133) + Rate Limit

The rules will be installed as:

User-Order 10

Component 4: Destination IP + Packet content filter (otype 2, offset 70, content-length 3 , content 0x111122) + Discard

Component 5: Destination IP + Packet content filter (otype 2, offset 70, content-length 3 , content 0x111133) + Rate Limit

Component 2: Destination IP + Packet content filter (otype 0, offset 50, content-length 3 , content 0x111122) + Discard

Component 3: Destination IP + Packet content filter (otype 2, offset 70, content-length 2 , content 0x1111) + Rate Limit

Component 1: Destination IP + Packet content filter (otype 0, offset 50, content-length 2 , content 0x1111) + Rate Limit

5. Operational Considerations

The Packet Content Filter is intended for controlled deployment scenarios, such as traffic filtering, traffic optimization, and DDoS mitigation based on known fixed packet-content patterns. Operators SHOULD deploy this filter only at controlled filtering locations, such as provider edge devices, traffic steering points, mitigation points, or other devices where the traffic impact and rollback procedures are well understood.

Operators SHOULD enable Packet Content Filter processing only on devices that support packet-content parsing and have sufficient filtering resources for the expected rule scale. Unsupported rules SHOULD be rejected or ignored locally according to local policy.

When encapsulation is present, such as MPLS, GRE, or other tunnels, the offset base can become ambiguous if matching is applied to the outer packet. Operators SHOULD apply matching to the decapsulated inner IP packet when applicable, or otherwise ensure that the offset base is unambiguous.

6. Scalability Considerations

Packet-content matching may consume limited implementation resources, such as UDF, ACL, or TCAM entries. Operators SHOULD limit Packet Content Filter rules to a small set of high-value entries, such as confirmed attack signatures, operationally validated filtering rules, or traffic optimization policies.

When FSv2 is used, rule ordering SHOULD be used to reduce the amount of traffic requiring packet-content inspection, for example by combining packet-content matching with more specific header-based conditions.

Operators SHOULD restrict the propagation scope of Packet Content Filter rules to avoid unnecessary inter-domain scale impact. Inter-domain propagation SHOULD be used only with explicit operational agreement and suitable policy control.

7. Security Considerations

This specification does not change the security properties of BGP itself. However, Packet Content Filter rules can affect traffic treatment and may cause packets to be dropped, redirected, rate-limited, or other actions according to local policy.

Operators MUST apply appropriate import policies, validation procedures, and authorization controls before accepting Packet Content Filter rules. Such rules SHOULD be accepted only from trusted BGP peers, and their propagation scope SHOULD be restricted by local policy.

To reduce false positives, Packet Content Filter rules SHOULD be combined with other FlowSpec match conditions, such as destination prefix, source prefix, protocol, port, TCP flags, or fragment-related conditions, when applicable.

Unsupported rules SHOULD be rejected or ignored locally according to local policy. Implementations and operators SHOULD apply update-rate limits and resource limits to avoid excessive control-plane load and preserve BGP stability.

8. IANA Considerations

IANA is requested to assign a new Type Value for the Packet Content Filter from the "Flow Spec Component Types" registry.

Type Value	Name	Reference
TBD	Packet Content filter	this document

For FSv2, a Packet Content Filter Component Type will be requested from the appropriate FSv2 Extended IP Filters component registry after that registry is defined.

9. Normative References

[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/rfc/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/rfc/rfc8956>.

[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/rfc/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/rfc/rfc8174>.

[I-D.ietf-idr-flowspec-v2] Hares, S., 3rd, D. E. E., Yadlapalli, C., and S. Maduschke, "BGP Flow Specification Version 2", April 2024, <https://datatracker.ietf.org/doc/draft-ietf-idr-flowspec-v2/04/>.

[I-D.hares-idr-fsv2-more-ip-filters] Hares, S. and N. Kao, "BGP Flow Specification Version 2 - More IP Filters", n.d., <https://datatracker.ietf.org/doc/draft-hares-idr-fsv2-more-ip-filters/>.

[I-D.khare-idr-bgp-flowspec-payload-match] Khare, A., BERGEON, P., Kestur, V., Jalil, L., and K. Kasavchenko, "BGP Flow Specification for Payload Matching", n.d., <https://datatracker.ietf.org/doc/draft-khare-idr-bgp-flowspec-payload-match/>.

Acknowledgements

We wish to thank Jeffrey Haas and Li Yang for their valuable comments and suggestions on this document. We also wish to thank Rui Xu and Yannan Hu for their contribution in the implementation and validation of the packet content filter software.

Authors' Addresses

Yong Cui
Tsinghua University
Beijing, 100084
China
Email: cuiyong@tsinghua.edu.cn
URI: http://www.cuiyong.net/

Yujia Gao
Zhongguancun Laboratory
Beijing, 100094
China
Phone: +86-185-1028-7458
Email: gaoyj@zgclab.edu.cn

Susan Hares
Hickory Hill Consulting
7453 Hickory Hill
Saline, Michigan 48176
United States of America
Email: shares@ndzh.com