IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: 17 September 2026

S. Hares
Hickory Hill Consulting
D. Eastlake
Independent
J. Dong
Huawei Technologies
C. Yadlapalli
ATT
S. Maduscke
Verizon
J. Haas
Juniper

16 March 2026

BGP Flow Specification Version 2 - for Basic IP
draft-ietf-idr-fsv2-ip-basic-04

Abstract

BGP flow specification version 1 (FSv1), defined in RFC 8955, RFC 8956, and RFC 9117 describes the distribution of traffic filter policy (traffic filters and actions) distributed via BGP. During the deployment of BGP FSv1 a number of issues were detected, so version 2 of the BGP flow specification (FSv2) protocol addresses these issues. In order to provide a clear demarcation between FSv1 and FSv2, a different NLRI encapsulates FSv2.

The IDR WG requires two implementation. Early feedback on implementations of FSv2 indicate that FSv2 has a correct design direction, but that breaking FSv2 into a progression of documents would aid deployment of the draft (basic, adding more filters, and adding more actions). This document specifies the basic FSv2 NLRI with user ordering of filters added to FSv1 IP Filters and FSv2 actions.

Copyright Notice

Table of Contents

1. Introduction

Version 2 of BGP flow specification was original defined in [I-D.ietf-idr-flowspec-v2] (BGP FSv2).

FSv2 is an update to BGP Flow specification version 1 (BGP FSv1). BGP FSv1 as defined in [RFC8955], [RFC8956], and [RFC9117] specified 2 SAFIs (133, 134) to be used with IPv4 AFI (AFI = 1) and IPv6 AFI (AFI=2).

The iinitial BGP FSv2 specification had the correct direction, but it contained than the early implementers desired. The imoplmenters desired a progression of documents with smaller incremental changes (basic FSv2, adding more filters, and adding more actions.)

This draft (FSv2 Basic) provides the basic FSv2 framework specification for transmitting user-ordered IP filters in the FSV2 NLRI and associating Flow Spec actions by transmitting the FLow Spec Extended Community(FS-EC) with the FSv2 NLRI. If a filter match links to a single FS-EC action, the single action succeeds or fails. If a filte rmatch links to mutiple actions, there is a potential for interactions. Section x.x discusses how to analyze the interaction by categories and solutions to issues with multiple FSv2-EC actions interacting. A complete solution requires the BGP Community Attribute see [I-D.ietf-idr-wide-bgp-communities]) with FSv2 Container defined in the [I-D.hares-idr-fsv2-more-ip-actions].

This document defines 2 new SAFIs (TBD1, TBD2) for FSv2 to be used with 5 AFIs (1, 2, 6, 25, and 31) to allow user-ordered lists of traffic match filters for user-ordered traffic match actions encoded in Extended Communities (FS-EC) or a newly defined BGP. FSv2 does not require combinations of FSv2 AFI/SAFIs to be implemented (10 combinations). An implementation is required to implement only 1 these combinations to be compliant. For example, a compliant implementation might only define the FSv2 NLRI for IPv4 for IP forwarding (AFI=1, SAFI=TBD1).

FSv1 and FSv2 use different AFI/SAFIs to send flow specification filters. Since BGP route selection is performed per AFI/SAFI, this approach can be termed “ships in the night” based on AFI/SAFI.

The remainder of section 1 provides background on why the FSv2 was necessary to fix problems with FSv1. Section 2 contains a Primer on FSv2. Section 3 contains the encoding rules for FSv2 and user-based encoding sent via BGP. Section 4 describes how to validate and order FSv2 NLRI. Sections 5-8 discusses scalability, optional security additions, security considerations, and IANA considerations.

1.1. Why Flow Specification v2

Modern IP routers have the capability to forward traffic and to classify, shape, rate limit, filter, or redirect packets based on administratively defined policies. These traffic policy mechanisms allow the operator to define match rules that operate on multiple fields within header of an IP data packet. The traffic policy allows actions to be taken upon a match to be associated with each match rule. These rules can be more widely defined as “event-condition-action” (ECA) rules where the event is always the reception of a packet.

BGP ([RFC4271]) flow specification as defined by [RFC8955], [RFC8956], [RFC9117] specifies the distribution of traffic filter policy (traffic filters and actions) via BGP to a mesh of BGP peers (IBGP and EBGP peers). The traffic filter policy is applied when packets are received on a router with the flow specification function turned on. The flow specification protocol defined in [RFC8955], [RFC8956], and [RFC9117] will be called BGP flow specification version 1 (BGP FSv1) in this draft.

Multiple deployed applications currently use BGP FSv1 to distribute traffic filter policy. These applications include: 1) mitigation of Denial of Service (DoS), 2) traffic filtering in BGP/MPLS VPNS, and 3) centralized traffic control for networks utilizing SDN control of router firewall functions, 4) classifiers for insertion in an SFC, and 5) filters for SRv6 (segment routing v6).

During the deployment of BGP flow specification v1, the following issues were detected:

* lack of a TLV encoding prevented extension of encodings (FSv1 uses TV (type-value),

* inability to allow users to define order for filtering rules,

* inability to allow users to define order for multiple actions, and

* lack of clear rules for multiple actions per filter match that provide default actions on the order of actions and what happens when an action fails.

Networks currently cope with these issues above by constraining deployments or using topology/deployment specific workaround.

FSv1 is a critical component of deployed applications. Therefore, this specification defines how FSv2 will interact with BGP peers that support either FSv2, FSv1, or FSv2 and FSv1. It is expected that a transition to FSv2 will occur over time as new applications require FSv2 features.

1.2. Definitions and Acronyms

AFI- Address Family Identifier

AS - Autonomous System

BGP Session ephemeral state - state which does not survive the loss of BGP peer session.

BGP Commmunity Path Attribute - BGP Community Path attribute with a FS TLV defined by [I-D.hares-idr-fsv2-more-ip-actions]

Configuration state - state which persist across a reboot of software module within a routing system or a reboot of a hardware routing device.

CPA - BGP Community Path Attribute

DDoS - Distributed Denial of Service.

Ephemeral state - state which does not survive the reboot of a software module, or a hardware reboot. Ephemeral state can be ephemeral configuration state or operational state.

FSv1 - Flow Specification version 1 [RFC8955] [RFC8956]

FSv2 - Flow Specification version 2 (this document)

FS - Flow Specification (either v1 or v2)

FS-CPA - Flow Specification Actions defined in Community Path Attribute

FS-EC - FS related Extended Community with FS actions

FSv1-EC - FSv1 Extended Community with FS Actions supported by FSv1

FSv2-EC - FSv2 Extended Community with FS Actions supported by FSv2

NETCONF - The Network Configuration Protocol [RFC6241].

RESTCONF - The RESTCONF configurati on Protocol [RFC8040]

RIB - Routing Information Base.

ROA - Route Origin Authentication [RFC9582]

RR - Route Reflector.

SAFI Subsequent Address Family Identifier

1.3. RFC 2119 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.

2. Flow Specification Version 2 Primer

A BGP Flow Specification (v1 or v2) is an n-tuple containing one or more match criteria that can be applied to IP traffic, traffic encapsulated in IP traffic or traffic associated with IP traffic. The exact traffic match depends on the FSv2 AFI/SAFI.

A given Flow Specification NLRI may be associated with a set of path attributes depending on the particular application, and these attributes within that set may or may not include reachability information (e.g., NEXT_HOP). The BGP path attributes associated with FSv2 may also contain Extended Community attributes [RFC4360] and Community Container Path attributes [I-D.ietf-idr-wide-bgp-communities] with FSv2 actions.

BGP processing treats the NLRI as a key to entries in AFI/SAFI BGP databases. Entries that are placed in the Loc-RIB are then associated with a given set of semantics which are application dependent. Standard BGP mechanisms such as update filtering by NLRI or by attributes such as AS_PATH or large communities apply to the BGP Flow Specification defined NLRI-types.

This FSv2 for basic IP forwarding specification only requires the use of Extended Communities to associate FS action with FSv2 filters found in FSv2 NLRI.

The FSv2 features of multiple actions with user ordering of actions or dependencies between actions requires the BGP Community Attribute [I-D.ietf-idr-wide-bgp-communities] with a FSv2 Component as defined in [I-D.hares-idr-fsv2-more-ip-actions].

Network operators can control the propagation of BGP routes by enabling or disabling the exchange of routes for a particular AFI/SAFI pair on a particular peering session. As such, the Flow Specification may be distributed to only a portion of the BGP infrastructure.

2.1. Flow Specification v1 (FSv1) SAFIs

The FSv1 NLRI defined in [RFC8955] and [RFC8956] include 13 match conditions encoded for the following AFI/SAFIs:

* IPv4 traffic: AFI:1, SAFI:133

* IPv6 Traffic: AFI:2, SAFI:133

* BGP/MPLS IPv4 VPN: AFI:1, SAFI: 134

* BGP/MPLS IPv6 VPN: AFI:2, SAFI: 134

Match conditions are ordered by component type in ascending order. If multiple component types filters exist, the ordering within a component type is defined by the component type. The FSv1 component format does not provide enough information to create a unique canonically sorted list for all implementations in all deployments.

The actions standardized for in [RFC8955] and [RFC8956] are:

* accept packet (default),

* traffic flow limitation by bytes (0x6),

* traffic-action (0x7),

* redirect traffic to VPN (0x8),

* mark traffic (0x9), and

* traffic flow rate limiting (12, 0xC)

An SFC action [RFC9015] defines a redirection of a data flow to an entry point into a specific SFP (Service Function Path)

While IDR has proposed other Extended Community Actions, no additional actions have completed the standardization process.

Additional proposals for FS Actions in Extended Communities exist, but these are still in the standardization process. As such, these new FS actions will fall under FSv2 rules for Extended Communities.

2.2. Transition to FSv2

This specification specification defines AFI/SAFI pairs to support Flow Spec for IPv4 IP IPv6, L2, IPv4 VPNs, IPv6 VPNs, L2VPN, SFC, and SFC VPN supports the components and actions for the following:

* IPv4 (AFI=1, SAFI=TBD1),

* IPv6 (AFI=2, SAFI=TBD1),

* L2 (AFI=6, SAFI=TDB1) [described in [I-D.ietf-idr-flowspec-l2vpn]),

* BGP/MPLS IPv4 VPN: (AFI=1, SAFI=TBD2),

* BGP/MPLS IPv6 VPN: (AFI=2, SAFI=TBD2),

* BGP/MPLS L2VPN (AFI=25, SAFI=TDB2) [described in [I-D.ietf-idr-flowspec-l2vpn]),

* SFC: (AFI=31, SAFI=TBD1),

* SFC VPN (AFI=31, SAFI=TBD2),

One question asked by developers is what AFI/SAFI is required for FSv2 IP Basic compliance. BGP negotiates support for each AFI/SAFI, so FSv2 IP Basic support for non-VPN could be as little as FSv2 for IPv4 forwarding (AFI/SAFI: 1/TBD1),

The IDR specification for L2 VPN traffic was specified in [I-D.ietf-idr-flowspec-l2vpn]. An IDR specification for tunneled traffic is in [I-D.ietf-idr-flowspec-nvo3]. Both of these drafts were targeted for FSv1, but the WG decided to require these to FSv2 TLV formats.

2.3. FSv2 Overview

FSv2 allows the user to order the flow specification rules and the actions associated with a rule. Each FSv2 rule may have one or more match conditions and one or more associated actions.

FSv2 operates in the ships-in-the night model with FSv1 so network operators can manipulate which the distribution of FSv2 and FSv1 using configuration parameters.

An FSv2 peer may receive BGP NLRI routes from a FSv1 peer or a BGP peer that does not support FSv1 or FSv2. The capabilities sent by a remote BGP peer indicate what FS supported by the AFI/SAFI (FSv1 NLRI or FSv2 NLRI or none).

The basic principles regarding ordering of flow specification filter rules are:

1) Rule-0 (zero) is defined to be 0/0 with the “permit-all” action.

2) FSv2 filter rules are ordered based on user-specified order.

- The user-specified order is carried in the FSv2 NLRI and a numerical lower value takes precedence over a numerically higher value. For rules received with the same order value, the filter rules are ordered by FSv2 component number and then by rules specified by the component.

3) Importing FSv1 into a combined FSv2 and FSv2 rule set requires starting FSv2 with Rule 1 and adding FSv1 rules are added after FSv2 rules.

- For example, BGP Peer A has FSv2 data base with 10 FSv2 rules (1-10). Suppose that, FSv1 user ordered FS are configured to start at 301 so 10 FSv1 rules are added at 301-310.

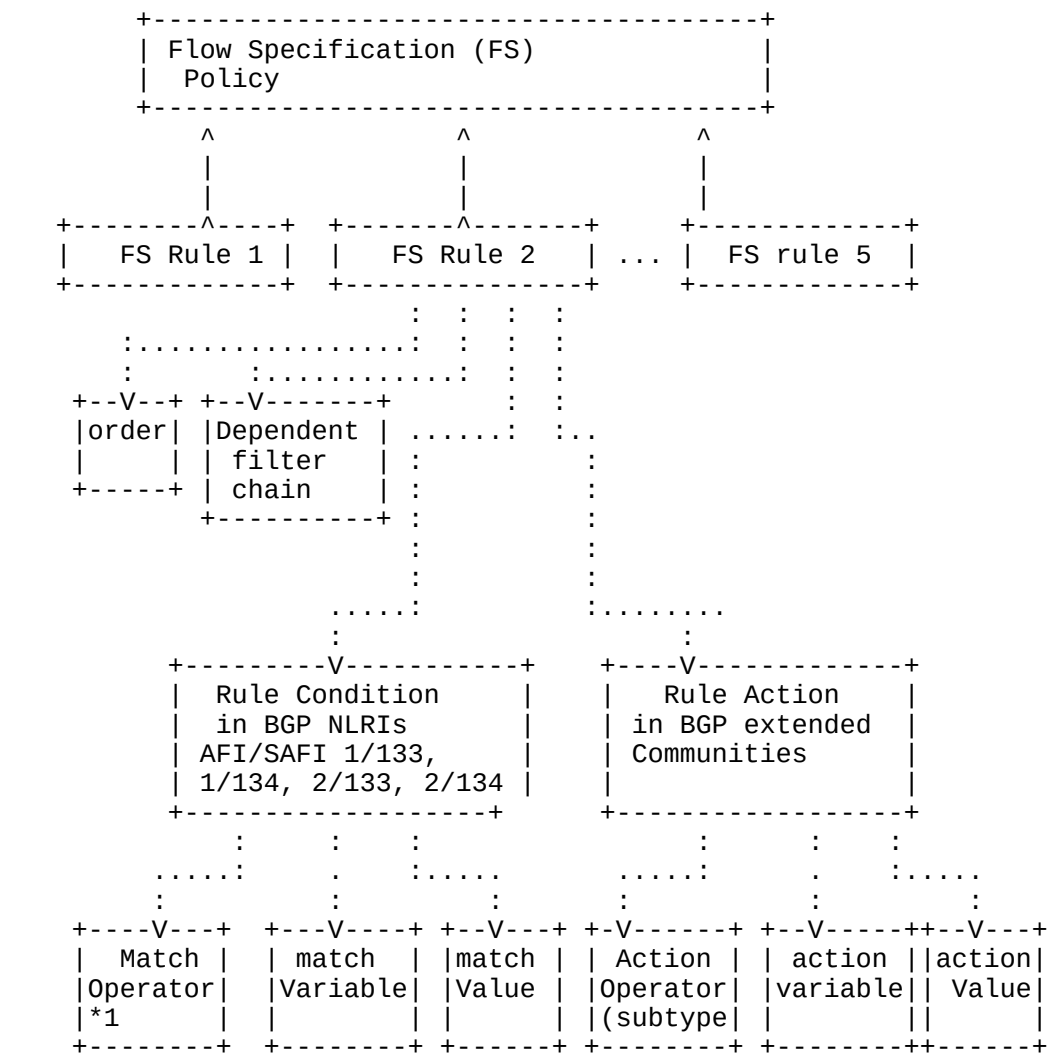
4) One or more action can be linked to a filter rule

- If a single action is linked, the action either succeeds or fails.

- if multiple actions are linked to a filter rule via FSv2 Extended Communities, the implementation MUST follow the FSv2-EC interaction rules in section x.x or have a local configuration knob to indicate local interaction rules. Implementations within a network SHOULD follow the same interaction rules

- If an implementation allows for FSv2 actions with user-ordering and Extended Community actions, the by default the Extended Community are ordered after the user-ordered actions.

Figure 1 shows a diagram of the FSv2 logical data structures with 5 rules for IP Basic functionality.



*1 match operator may be complex.

Figure 2-1: BGP Flow Specification v1 Policy

3. FSv2 NLRI Formats and Actions

3.1. FSv2 NLRI Format

The BGP FSv2 supports NRLI with the format for AFIs for IPv4 (AFI = 1), IPv6 (AFI = 2), L2 (AFI = 6), L2VPN (AFI=25), and SFC (AFI=31) with SAFIs TBD1 (Flow Spec) and TBD2 (Flow Spec for VPNs) to support transmission of the flow specification which supports user ordering of traffic filters and actions for IP traffic and IP VPN traffic.

A compliant FSv2 implementation only has to implement one AFI/SAFI pair out of the full list of NRLIs. For example, a compliant FSv2 implementation could only implement IPv4 FSv2 (AFI=1, SAFI=TBD1).

This NLRI information is encoded using MP_REACH_NLRI and MP_UNREACH_NLRI attributes defined in [RFC4760]. When advertising FSv2 NLRI, the length of the Next-Hop Network Address MUST be set to 0. Upon reception, the Network Address in the Next-Hop field MUST be ignored.

Implementations wishing to exchange flow specification rules MUST use BGP’s Capability Advertisement facility to exchange the Multiprotocol Extension Capability Code (Code 1) as defined in [RFC4760], and indicate a capability for FSv1, FSv2 (Code TBD3), or both.

The AFI/SAFI NLRI for BGP Flow Specification version 2 (FSv2) has the format:

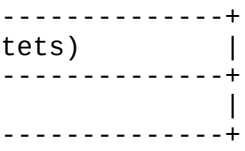


Figure 3-1 - NLRI format

where:

* NLRI length: length of field including all SubTLVs in octets.

* TLV+ - indicates the repetition of the TLV field

Each each TLV has the Format:

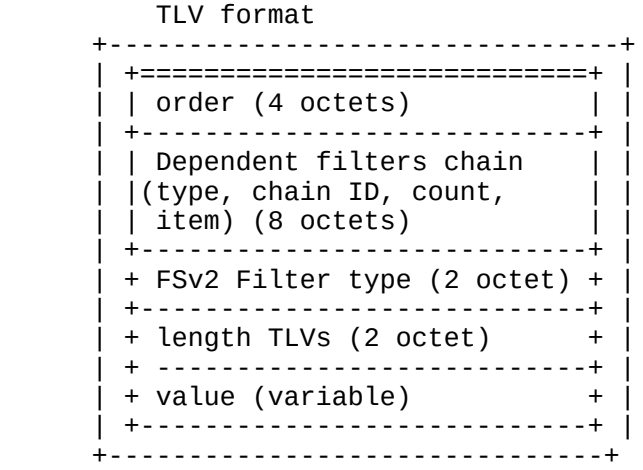


Figure 3-2 - TLV format within FSv2 NLRI

where:

* order: flow-specification global rule order number (4 octets).

* Dependent Filters Chain: 8 octets for identifying a chain of FSv2 filters that must be deployed at the same time.

  Why needed in FSv2: Flow specification filters distributed in BGP UPDATE packets may be broken into multiple packets. In FSv2, the dependent filter ID allows the filter chains to be identified across all user-defined or default filters. The rules can be installed from BGP into the firewall after all filters have been installed.

  For basic FSV2: This field is required to be set to all zero, and ignored upon reception.

  For future FSV2: Future specifications will specify the use of this field, and future specifications will continue to ignore the field if the value is all zeros.

* FSv2 Filter type: contains a type for FSv2 TLV format of the NRLI (2 octets). This type specifies support for a set of components. [Editor's note: We need to decide between Option 2 or Option 1. Option 2 makes provides a better default ordering for frame/packet filters.

  Option 1 ascending order (draft-ietf-idr-fsv2-ip-basic-02)

0 - reserved,

1 - IP Basic Filter rules

2 - IP Extended Filter Rules

3 - MPLS Filter Rules

4 - L2 traffic rules

5 - SFC Traffic rules

6 - Tunneled traffic

Option 2 ordering by frame/packet (new)

0 - reserved,

50 - L2 Traffic fules

100 - MPLS traffic rules

150 - SFC Traffic rules

200 - Tunneled traffic

256 - IP Basic Filter Rules (bit 1 of high bit)

280 - IP Extended Filter Rules

* length-TLV: is the length of the value part of the Sub-TLV,

* value: value depends on the type of FSv2 Filter type.

FSv2 implementations MUST pass valid filter TLVs even if the implementation does not support these installation of these a particular type of filter rules.

This specification only defines operation of the IP Basic Filter Rules that all FSv2 must support.

3.1.1. Ordering of TLVs within the FSv2 NLRI

For ease of processing, the ordering within the FSv2 NLRI MUST be by order number. Within an order value (e.g. 20), the filters MUST group the filters by the same filter type (e.g. IP Basic filter rules), and order the groups of filters by filter type (ascending values).

The order within a filter type (e.g. IP Basic Filters) MUST be by the component type. If multiple components of the same type exist, the component ordering is specified by the component definition (as in FSv1).

3.1.2. Partial Deployments

Partial deployments can occur for two reasons:

* Only a portion of the nodes in a network with FSv2 support installing new FSv2 Filter types with new FSv2 components. Other nodes (such as RRs), check the syntax, but do not handle the semantic meaning.

* During upgrades, a portion of the nodes know about a new Filter type with the components, but other nodes do not.

Editor: Are there others?

3.2. FSv2 Basic IP Filters

3.2.1. Operators for comparison

3.2.1.1. Numeric Operator (numeric_op)

This operator is encoded as shown in Figure 3-3.

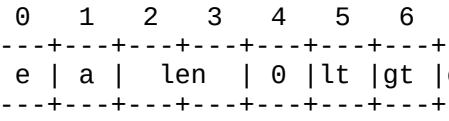


Figure 3-3: Numeric Operator (numeric_op)

e (end-of-list bit): Set in the last {op, value} pair in the list

a (AND bit): If unset, the result of the previous {op, value} pair is logically ORed with the current one. If set, the operation is a logical AND. In the first operator octet of a sequence, it MUST be encoded as unset and MUST be treated as always unset on decoding. The AND operator has higher priority than OR for the purposes of evaluating logical expressions.

len (length): The length of the value field for this operator given as (1 << len). This encodes 1 (len=00), 2 (len=01), 4 (len=10), and 8 (len=11) octets.

0 MUST be set to 0 on NLRI encoding and MUST be ignored during

decoding

lt less-than comparison between data and value

gt: greater-than comparison between data and value

eq: equality between data and value

The bits lt, gt, and eq can be combined to produce common relational operators, such as "less or equal", "greater or equal", and "not equal to", as shown in Table 3-1.

lt	gt	eq	Resulting operation
0	0	0	false (independent of the value)
0	0	1	== (equal)
0	1	0	> (greater than)
0	1	1	<= (greater than or equal)
1	0	0	< (less than)
1	0	1	<= (less than or equal)
1	1	0	!= (not equal value)
1	1	1	true (independent of the value)

Table 3-1: Comparison Operation Combinations

3.2.1.2. Bitmask Operator (bitmask_op)

This operator is encoded as shown in Figure 3-4.

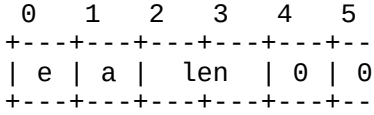


Figure 3-4 Bitmask Operator (bitmask_op)

Where:

e, a, len (end-of-list bit, AND bit, and length field): Most significant nibble; defined in the Numeric Operator format in section 3-x.

not (NOT bit): If set, logical negation of operation.

m (Match bit): If set, this is a bitwise match operation defined as "(data AND value) == value"; if unset, (data AND value) evaluates to TRUE if any of the bits in the value mask are set in the data.

0 (all 0 bits): MUST be set to 0 on NLRI encoding and MUST be ignored during decoding

3.2.2. IP Basic Filters (Filter type=TBD)

The format of the IP Basic TLV value field is shown in Figure 3-5. The IP header for the VPN case is specified in section 3.5.

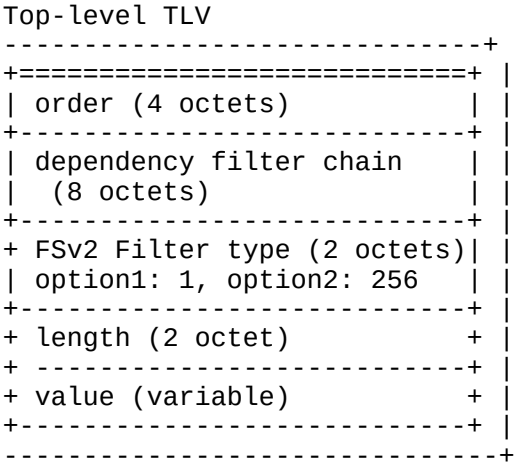


Figure 3-5 NLRI format for FSv2 IP Filter Type

Where:

order - is an 4 octet field with a value 1-N. The value 0 (zero) is invalid, and the TLV should be "treated-as-withdrawl".

dependency filter chain - is an 8 octet field which must be all zero for the IP Basic Filter rules.

length - is a 2 octet field indicating the length of the value field.

value - is a variable field comprised of a sequence of component TLVs:

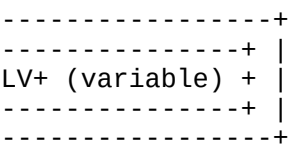


Figure 3-6 Value Field

Where the Component TLVs are:

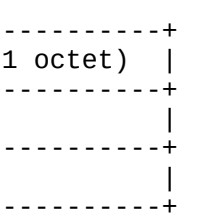


Figure 3-7 IP header Component TLVs

Where:

- Component type: component values are defined in the “Flow Specification Component types” registry for IPv4 and IPv6 by [RFC8955], [RFC8956], and [I-D.ietf-idr-flowspec-srv6]

- length: length of SubTLV (varies depending on the component type). If the length of the component types does match the valid defined length(s) for the component, the component type is ignored and the Filter type TLV is "treated-as-withdrawl".

- value: dependent on component type.

Many of the components use the operators [numeric_op] and [bitmask_op] defined in [RFC8955]

The list of valid SubTLV types appears in Table 3-2 for filter type of IP Filters (type=1). Other filters beyond these filters may be defined other filter types (e.g. IP Extended Filters).

Table 3-2 FSv2 IP Basic TLV Components

Sub-TLV	Definition
0 -	Reserved
1 -	IP Destination prefix
2 -	IP Source prefix
3	IPv4 Protocol / IPv6 Upper Layer Protocol
4	Port
5	Destination Port
6	Source Port
7	ICMPv4 type / ICMPv6 type
8	ICMPv4 code / ICPv6 code
9	TCP Flags
10	Packet length
11	DSCP
12	Fragment
13	Flow Label
14-255	Reserved

3.2.3. Ordering within the IP Basic Filter TLVs

The ordering of components within the value field of the IP Basic TLV follows the FSv1 rules. The following is a restatement of FSv1 rules in FSv2 terms.

* 1) order by component types (1-13).

* (2) If the components are the same, then the value fields are compared using mechanisms defined in [RFC8955] and [RFC8956] and MUST be in ascending order. NLRIs having component TLVs which do not follow the above ordering rules MUST be considered as malformed by a BGP FSv2 propagator. This rule prevents any ambiguities that arise from the multiple copies of the same NLRI from multiple BGP FSv2 propagators. A BGP implementation SHOULD treat such malformed NLRIs as "treat-as-withdrawl". [RFC7606].

See [RFC8955], [RFC8956], and for details on per component ordering.

3.2.4. FSv2 Components for IP Basic TLVs

3.2.4.1. IP Destination Prefix (type = 1)

IPv4 Name: IP Destination Prefix (reference: [RFC8955])

IPv6 Name: IPv6 Destination Prefix (reference: [RFC8956])

IPv4 length: Prefix length in bits

IPv4 value: IPv4 Prefix (variable length)

IPv6 length: length of value

IPv6 value: [offset (1 octet)] [pattern (variable)] [padding(variable)]

If IPv6 length = 0 and offset = 0, then component matches every address. Otherwise, length must be offset "less than" length "less than" 129 or component is malformed.

3.2.4.2. IP Source Prefix (type = 2)

IPv4 Name: IP Source Prefix (reference: [RFC8955])

IPv6 Name: IPv6 Source Prefix (reference: [RFC8956])

IPv4 length: Prefix length in bits

IPv4 value: Source IPv4 Prefix (variable length)

IPv6 length: length of value

IPv6 value: [offset (1 octet)] [pattern (variable)][padding(variable)]

If IPv6 length = 0 and offset = 0, then component matches every address. Otherwise, length must be offset < length < 129 or component is malformed.

3.2.4.3. IP Protocol (type = 3)

IPv4 Name: IP Protocol IP Source Prefix (reference: [RFC8955])

IPv6 Name: IPv6 Upper-Layer Protocol: (reference: [RFC8956])

IPv4 length: variable

IPv4 value: [numeric_op, value]+

IPv6 length: variable

IPv6 value: [numeric_op, value]+

where the value following each numeric_op is a single octet.

3.2.4.4. Port (type = 4)

IPv4/IPv6 Name: Port (reference: [RFC8955]), [RFC8956])

Filter defines: a set of port values to match either destination port or source port.

IPv4 length: variable

IPv4 value: [numeric_op, value]+

IPv6 length: variable

IPv6 value: [numeric_op, value]+

where the value following each numeric_op is a single octet.

Note-1: (from FSV1) In the presence of the port component (destination or source port), only a TCP (port 6) or UDP (port 17) packet can match the entire flow specification. If the packet is fragmented and this is not the first fragment, then the system may not be able to find the header. At this point, the FSv2 filter may fail to detect the correct flow. Similarly, if other IP options or the encapsulating security payload (ESP) is present, then the node may not be able to describe the transport header and the FSv2 filter may fail to detect the flow.

The restriction in note-1 comes from the inheritance of the FSv1 filter component for port. If better resolution is desired, a new FSv2 filter should be defined.

Note-2: FSv2 component only matches the first upper layer protocol value.

3.2.4.5. Destination Port (type = 5)

IPv4/IPv6 Name: Destination Port (reference: [RFC8955]), [RFC8956])

Filter defines: a list of match filters for destination port for TCP or UDP within a received packet

Length: variable

Component Value format: [numeric_op, value]+

3.2.4.6. Source Port (type = 6)

IPv4/IPv6 Name: Source Port (reference: [RFC8955]), [RFC8956])

Filter defines: a list of match filters for source port for TCP or UDP within a received packet

IPv4/IPv6 length: variable

IPv4/Ipv6 value: [numeric_op, value]+

3.2.4.7. ICMP Type (type = 7)

IPv4: ICMP Type (reference: [RFC8955])

Filter defines: Defines: a list of match criteria for ICMPv4 type

IPv6: ICMPv6 Type (reference: [RFC8956])

Filter defines: a list of match criteria for ICMPv6 type.

IPv4/IPv6 length: variable

IPv4/IPv6 value: [numeric_op, value]+

3.2.4.8. ICMP Code (type = 8)

IPv4: ICMP Type (reference: [RFC8955])

Filter defines: a list of match criteria for ICMPv4 code.

IPv6: ICMPv6 Type (reference: [RFC8956])

Filter defines: a list of match criteria for ICMPv6 code.

IPv4/IPv6 length: variable

IPv4/IPv6 value: [numeric_op, value]+

3.2.4.9. TCP Flags (type = 9)

IPv4/IPv6: TCP Flags Code (reference: [RFC8955])

Filter defines: a list of match criteria for TCP Control bits

IPv4/IPv6 length: variable

IPv4/IPv6 value: [bitmask_op, value]+

Note: a 2 octets bitmask match is always used for TCP-Flags

3.2.4.10. Packet length (type = 10 (0x0A))

IPv4/IPv6: Packet Length (reference: [RFC8955], [RFC8956])

Filter defines: a list of match criteria for length of packet (excluding L2 header but including IP header).

IPv4/IPv6 length: variable

IPv4/IPv6 value: [numeric_op, value]+

Note:[RFC8955] uses either 1 or 2 octet values.

3.2.4.11. DSCP (Differentiaed Services Code Point)(type = 11 (0x0B))

IPv4/IPv6: DSCP Code (reference: [RFC8955], [RFC8956])

Filter defines: a list of match criteria for DSCP code values to match the 6-bit DSCP field.

IPv4/IPv6 length: variable

IPv4/IPv6 value: [numeric_op, value]+

Note: This component uses the Numeric Operator (numeric_op) described in [RFC8955] in section 4.2.1.1. Type 11 component values MUST be encoded as single octet (numeric_op len=00).

The six least significant bits contain the DSCP value. All other bits SHOULD be treated as 0.

3.2.4.12. Fragment (type = 12 (0x0C))

IPv4/IPv6: Fragment (reference: [RFC8955], [RFC8956])

Filter defines: a list of match criteria for specific IP fragments.

Length: variable

Component Value format: [bitmask_op, value]+

Bitmask values are:

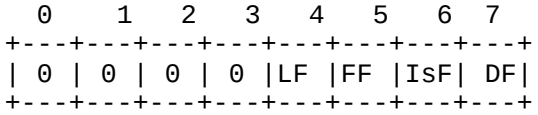


Figure 3-8

Where:

DF (don’ t fragment): match If IP header flags bit 1 (DF) is 1.

IsF(is a fragment other than first: match if IP header fragment offset is not 0.

FF (First Fragment): Match if [RFC0791] IP Header Fragment offset is zero and Flags Bit-2 (MF) is 1.

LF (last Fragment): Match if [RFC0791] IP header Fragment is not 0 And Flags bit-2 (MF) is 0

0: MUST be sent in NLRI encoding as 0, and MUST be ignored during reception.

3.2.4.13. Flow Label(type = 13 (0xOD))

IPv4/IPv6: Fragment (reference: [RFC8956])

Filter defines: a list of match criteria for 20-bit Flow Label in the IPv6 header field.

Length: variable

Component Value format: [numeric_op, value]+

3.3. FSv2 Actions for FSv2 IP Basic

The IP Basic FSv2 allows FS actions to be sent in an Extended Community (FSv2-EC) for IPv4 and IPv6. The Extended Community encodes the Flow Specification actions in the Extended IPv4 Community format [RFC4360] or in the extended IPv6 Community format [RFC5701].

A Flow Spec filter match maybe linked to only one Flow Spec action. For these deplyments, the details on multiple actions per flow spec filter match in this section can be ignored.

When an operator defines multiple Flow Spec Extended Community (FS-EC) Actions for single filter match, it is possible that some FS-EC actions can interact If one sorts the FS-EC actions into categories, the interactions can be minimized. One example of such interactions is if the flow is redirect to a VRF and to a VPN. Another way two actions interact is if an action fails. For example, if set DSCP action fails prior to redirect to a VPN may have undesired data flow patterns.

This section defines the FS-EC actions, categories of FS-EC actions, and ways to minimize interactions between FS-EC actions. Section 3.3.1 describes the existing FS-EC action formats. Section 3.3.2 describes the interaction between FS-EC action, and categories of actions, and the way to minimizes interactions. Should the BGP to distribute information about the configured interactions for sequences of multiple actions, Section 3.3 defines an optional FS-EC to pass information ordering of categories (user/this standard) and failure action (stop or best effort).

Note that FSv2 implementations that only associate 1 FSv2-EC per filter match do not need the FSv2-EC.

3.3.1. FSv2 Extended Community Actions

The FSv2 IP Basic uses FSv1 actions and defines for one one additional optional FSv2 specific FS-EC. This one optional action is the Action Chain Ordering (ACO) Extended Community (ACO-EC) which can pass around defaults currently only available by configuration in FSv1.

3.3.1.1. FSv2 Actions in Extended Community for IPv4

The format of the Extended Community for IPv4 defined in [RFC4360] is shown in Figure 3-9 with 2 octet type that is split into a high byte and low byte. The format of the IPv4 Extended Community is shown in Figure 3-10.

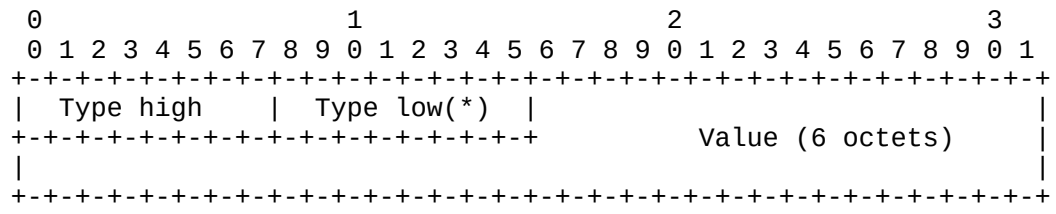


Figure 3-9

Table 3-3
FSv1 Transitive Extended Communities for IP Basic
High-Low byte of Transitive FS-EC

H-L	FSv1 Description	Short-ID	FS document
0x01-0C	Transitive IPv4	RDIPv4	RDIP
0x07-02	FSv1 for an Interface set	TAIS	ifset
0x09-xx	Redirect to Indirection ID	RGID	RGID
0x0b-00	SFC Reserved	SFC-R	RFC9015
0x0b-01	SFVC SFIR POOL Identifier	SFIR-PI	RFC9015
0x0b-02	SFC MPLS label stack Swapping or stacking labels	SFC-MPLS	RFC9015
0x80-06	Traffic rate limit by bytes	TRB	RFC8955
0x80-07	Traffic Action (sample, terminal)	TA	RFC8955
0x80-08	Redirection to VRF (2 AS form)	RDIP	RFC8955
0x80-09	Traffic mark DSCP	TM	RFC8955
0x80-0C	Traffic rate limit by packets	TRP	RFC8955
0x81-08	Redirect to VPN (IPv4 form)	RDIP	RFC8955
0x81-08	Redirect to VPN (4 AS form)	RDIP	RFC8955

Note the Short ID is simply a quick way for this document to reference a particular action.

References:

ifset: [I-D.ietf-idr-flowspec-interfaceset]

RDIP: [I-D.ietf-idr-flowspec-redirect-ip]

RGID: [I-D.ietf-idr-flowspec-path-redirect]

RFC9015: [RFC9015]

RFC8955: [RFC8955]

3.3.1.2. Flow Specification Actions in IPv6 forms

The Transitive IPv6-Address-Specific Extended Community encodes the Flow Specification actions in the Extended Community format specified in [RFC5701] shown in Figure 3-10. Table 3-3 lists the 4 octet format for high-byte and low-byte. Note that there are two allocations for redirect from IPv6.

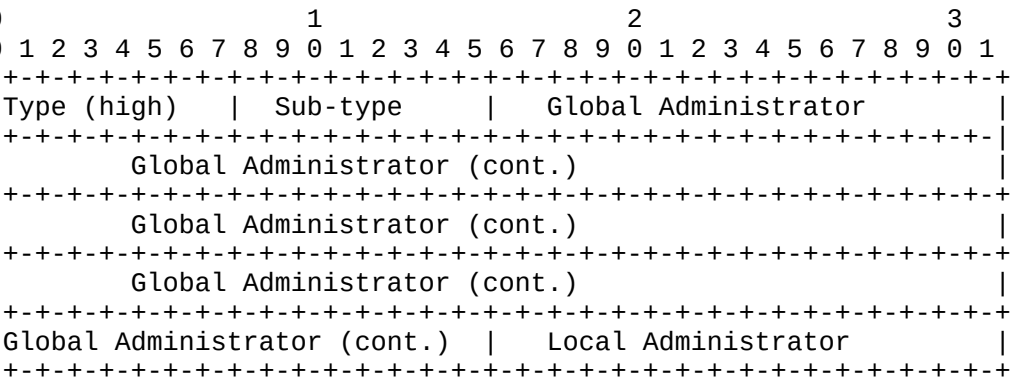


Figure 3-10

The 20 octets of value are given in the following format:
Global Administrator: IPv6 address assigned by Internet Registry
Local Administrator: 2 bytes of Local Administrator

Table 3-4
Transitive IPv6-Address-Specific Extended Community Types

H-L byte	FSv1 Description	Short-ID	FS document
0x0000	Unassigned		
0x0001	Unassigned		
0x0002	Route Target	RT	No
0x0003	Route origin	ROrg	No (deprecated)
0x0004	OSPFv3 Route Attribute	OSPFv3	No (deprecated)
0x0005	FIT Tail Community	IFITv6	No
0x0006	Link Bandwidth	LBW	No
0x0008	Unassigned		
0x0009	Unassigned		
0x000A	Unassigned		
0x000B	VRF Route Import	VRP-I	No
0x000C	FS Redirect to IPv6	RDIPv6C	Yes: RDIP
0x000D	FS Redirect to IPv6	RDIPv6	Yes: RFC8956
0x000E	Unassigned		
Ox000F	Unassigned		
0x0010	Cisco VPN distinguisher	Cisco-VPN	No
0x0011	UUID-based Route Target	UUID-RT	No
0x0012	Inter-ARea P2MP S-NH	PM2P-NH	No
0x0013	Unassigned		
0x0014	VRF Recursive NH	VRF-RNH	no
0x0015	RT-derived-EC	RT-EC	no

3.3.2. Interactions between Flow Spec Extended Community Actions (FS-EC)

Interactions between FS-EC only occur if a filter match has multiple actions. If multiple FS-EC actions are listed on a filter, these can interact within categories or if a previous FS-EC action fails. Consider three cases of multiple actions as you read this sectio.

Case 1: a filter match has three actions:

1. set a DSCP value in the IP Packet,

2. Sample the traffic (for spam checking), and

3. Redirect the traffic to a VRF.

Case 2: a filter match has two actions.

1. redirect to VRF,

2. TAIS Terminate FS filter processing

Case 3: a Filter match has two actions

1. redirect to VRF, and

2. redirect to SFC

3.3.2.1. Interactions Between Multiple Actions by Categories

This section considers interactions between multiple successful FSv2 Actions.

FS-EC actions fall into the following categories:

1. limitation on filters or actions (interface-set or local configuration)

2. rate limiting (bytes (TRB) and packets (TRP)),

3. Set DSCP value in IP packet

4. Sample packet (TAIS - sample)

5. redirect to IP paths (to VRF, to VPN, to Indirection-ID, to SFC path)

6. Terminate action processing (TAIS terminate)

If multiple actions with only one action per category, then there is little interaction between successful actions. If multiple actions have more than one action per category, then successful actions can interact. For example, the multiple actions in Case 1 and Case 2 described above only have one type of action per category. In contrast, Case 3 has two interactions in the redirect category.

For multiple actions passed in the FS-EC, this logical categorization allow the implementation to reject a set of multiple actions if some of the actions interact in a set. Using the three case examples, Case 1 and Case 2 FS-EC would be allowed and Case 3 FS-AC would be rejected.

FSv2 actions passed in a BGP Community Attribute can provide ordering of actions, dependencies or signal which actions are valid within a category (see [I-D.hares-idr-fsv2-more-ip-actions]). However, these features are beyond the Basic FSv2 for IP forwarding and out of scope for this specification.

3.3.2.2. Failure of an FS-EC Action

If multiple FS-EC action is attached to a Flow Spec filter rule and one of the actions fails there are three potential options:

Option 1. Stop processing additional filters and (optionally) signal failure to the management process,

Option 2. Continue on processing in "best effort" for the next filters.

Option 3. Decide between 1 and 2 based on dependencies between filters and actions

Option 1 and 2 can be signaled by configuration within a Flow Specification implementation. Option 3 requires the encoding dependency lists in ordered filters and ordered actions. The FSv2 NLRI format has a field to carry filter dependency information, but these funtions are beyond the FSv2 Basic IP functions and out of scope for this specification.

Consider Case 1 where the set DSCP value in IP field fails to occur. Option 1, would be to stop processing and not do the other two actions. Option 2, would be to continue prcoessing and do the other two actions.

Currently, for FSv1 local configuration determines what happens if one of the actions fails within a set of multiple actions attached to a filter rule.

One option for FSv2 is to pass another FS-EC indicating what the originator expects will happen upon failure of an action.

3.3.2.3. Action Chain Ordering FSv2-EC (ACO) (optional)

Summary: >This optional FSv2-EC passes information on what the BGP peer originating the FSv2-EC expects will happen with multiple actions attached to a single filter.

Description: The BGP peer originating multiple FSv2 FS-EC actions attached to FSv2 NLRI (filters) may attach the Action Chain Ordering (ACO) FS-EC to inform BGP Peers receiving the FSv2 information how the originating pair expects action interactions and actions failures will be handled. Two fields are encoded in this FS-EC:

AC-interaction - What happens if two actions are specified in a category, and

AC-Failure - what happens if an action with multiple action set fails.

Encoding: The Generic Transitive encoding is shown in figure 3-11 with the field definitions below.

Generic Transitive Extended Community (IPv4)

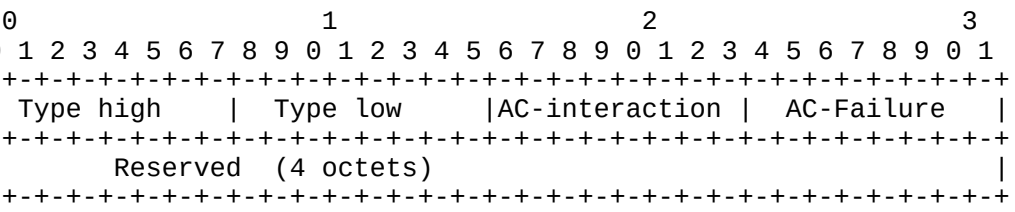


Figure 3-11

where:

Type high: This 1 octet field has a value of 0x80 For the Generic Transitive EC.

Type low: This one octet field identifies the ACO-Action. The value is TBD4.

ACO Dependency This field indicates whether the FS-EC category order is the pre-defined order or an implementation specific order.

0 (default): do not install actions with two actions per category.

1 (local config): allow under local configuration

AC-failure 1 octet byte that determines the action on failure. Actions may succeed or fail and an Action chain must deal with it. The default value stored for an action chain that does not have this action chain is “stop on failure” . where AC-Failure types are:

0x00 stop on failure of an action

0x01 continue on failure of an action,,

Reserved - Reserved for future use. Must be set to all zeros, and ignored upon reception.

4. Validation and Ordering of NLRI

4.1. Validation of FSv2 NLRI

The validation of FSv2 NLRI adheres to the combination of rules for general BGP FSv1 NLRI found in [RFC8955], [RFC8956], [RFC9117]. These FSv1 rules are sufficient for FSv2 for IP traffic.

Specific additions have been defined for IP Filters used for guiding IP traffic into Service Function Service Function Pathways SFC NLRI in [RFC9015], or validation of L2VPN FS NLRI (see [I-D.ietf-idr-flowspec-l2vpn]). These additions are not required for the FSv2 for IP Basic functions. Therefore, FSv2 NLRI validation for Basic IP uses the same rules as FSv1.

To provide clarity, the full validation process for flow specification routes (FSv1 or FSv2) for all AFI/SAFIs is described below in section x.x rather than simply referring to the relevant portions of these RFCs. Validation only occurs after BGP UPDATE message reception and the FSv2 NLRI and the path attributes relating to FSv2 (Extended community and Wide Community) have been determined to be well-formed. Any MALFORMED FSv2 NRLI is handled as a “session reset” [RFC7606].

4.1.1. Validation of FS NLRI (FSv1 or FSv2)

Flow specifications received from a BGP peer that are accepted in the respective Adj-RIB-In are used as input to the route selection process. Although the forwarding attributes of the two routes for tbe same prefix may be the same, BGP is still required to perform its path selection algorithm in order to select the correct set of attributes to advertise.

The first step of the BGP Route selection procedure (section 9.1.2 of [RFC4271] is to exclude from the selection procedure routes that are considered unfeasible. In the context of IP routing information, this is used to validate that the NEXT_HOP Attribute of a given route is resolvable.

The concept can be extended in the case of the Flow Specification NLRI to allow other validation procedures.

The FSv2 validation process validates the FSv2 NLRI with following unicast routes received over the same AFI (1 or 2) but different SAFIs:

FSv2 received over SAFI=TBD1 FSv1 received over SAFI=133 are be

validated against SAFI=1. Similarly, FSv2 routes received over SAFI=TBD1 will be validated against SAFI=1.

FSv2 received over SAFI=TBD2 FSv1 received over SAFI=134 are validated against SAFI=128, and FSv2 received over SAFI=TBD2 will be validated against SAFI=128.

FSv2 received with AFI = 31 The FSV2 routes received with (AFI=31, SAFI=TBD1) will be validated against SAFI=1. The FSv2 received with (AFI=31, SAFI=TBD2) will be validated against SAFI=128.

FSv2 L2 routes passed in (AFI=6, SAFI=TBD1) and L2VPN routes passed in (AFI=25, SAFI=TBD2). FSv2 L2 routes - validate (AFI=6, SAFI=TBD1) against (AFI=1, SAFI=1).

FSv2 L2VPN routes - validate (AFI=256, SAFI=TBD2) against (AFI=1, SAFI=128>

This is similar to FSv1. - The FSv1 L2 validated L2 routes passed in (AFI=6, SAFI=133) against (AFI=1, SAFI=1) and the L2VPN routes (AFI=25, SAFI=134) are validated against (AFI=1, SAFI=128).

In the absence of explicit configuration, a Flow specification NLRI (FSv1 or FSv2) MUST be validated such that it is considered feasible if and only if all of the conditions are true:

a) A destination prefix component is embedded in the Flow Specification,

b) One of the following conditions holds true:

- 1. The originator of the Flow Specification matches the originator of the best-match unicast route for the destination prefix embedded in the flow specification (this is the unicast route with the longest possible prefix length covering the destination prefix embedded in the flow specification).

- 2. The AS_PATH attribute of the flow specification is empty or contains only an AS_CONFED_SEQUENCE segment [RFC5065].

  - 2a.This condition should be enabled by default.

  - 2b.This condition may be disabled by explicit configuration on a BGP Speaker,

- o 2c.As an extension to this rule, a given non-empty AS_PATH (besides AS_CONFED_SEQUENCE segments) MAY be permitted by policy].

c) There are no “more-specific” unicast routes when compared with the flow destination prefix that have been received from a different neighbor AS than the best-match unicast route, which has been determined in rule b.

However, part of rule a may be relaxed by explicit configuration, permitting Flow Specifications that include no destination prefix component. If such is the case, rules b and c are moot and MUST be disregarded.

By “originator” of a BGP route, we mean either the address of the originator in the ORIGINATOR_ID Attribute [RFC4456] or the source address of the BGP peer, if this path attribute is not present.

A BGP implementation MUST enforce that the AS in the left-most position of the AS_PATH attribute of a Flow Specification Route (FSv1 or FSv2) received via the Exterior Border Gateway Protocol (eBGP) matches the AS in the left-most position of the AS_PATH attribute of the best-match unicast route for the destination prefix embedded in the Flow Specification (FSv1 or FSv2) NLRI.

The best-match unicast route may change over time independently of the Flow Specification NLRI (FSv1 or FSv2). Therefore, a revalidation of the Flow Specification MUST be performed whenever unicast routes change. Revalidation is defined as retesting rules a to c as described above.

4.1.2. Validation of Flow Specification Actions for FSv2 for IP Basic

A match on a Flow Specification (FS) filters is linked to one or more FS action set by an Extended Communities (FS-EC) for FSv2 for IP Basic functions.

Validation of FS-EC action begins with validating the syntax of the Extended Communities attributes attached to FS NLRI in UPDATE packet. Since FSv1 and FSv2 operate on different NLRIs (AFI/SAFI sets), a single FS-EC action can apply to both FSv1 and FSv2 filters. If the FS-EC is not syntactically correct, the FS-EC community causes NLRI and FS-EC to be treated as withdrawal.

If the FS-EC is syntacically correct, then the FS-EC check to determine if this node can perform this action. If not, the FS-EC is stored for transmittal to other nodes, but cannot be used in this node.

If multiple syntactically correct actions that can be performed on are linked to the filtering rules defined in the NLRI in UPDATE packet, then the list of multiple actions are check for conflicts within a category. If conflicts exist within a multiple action set attached to a FSv2 filter, then the default case is to ignore the action set for installation in the node. Optionally, if the ACO FS-EC may indicate if the BGP peer originating the FSv2 filter + action expects this "ignoring" of the action or specifical local configuration.

An example of local configuration might be if rate limiting by byte and by packet are specified, the local configuration might allow both to be enacted in the hardware.

4.1.3. Failure of FS-EC action when processing a filter flow

If one action in the ordered list fails for a traffic flow, the local node may be able to halt processing of the for. For example, if a DSCP value set and forwarding to VPN is specified AND the DSCP fails, the forwarding logic may allow the forwarding to the VPN to not occur.

FSv1-EC current control the failure action by configuration and/or implementation defaults.

The optional ACO FSv2-EC can inform the BGP receiving the FSv2 information how the originator expects failures within the multiple actions in an action set will occur. The ACO FSv2-EC is optional.

FSv2 Implementations MAY wish to log the action failures encountered by FS actions (FSv1 or FSv2).

4.1.4. Error handling and Validation

The following two error handling rules must be followed by all BGP speakers which support FSv2:

* FSv2 NLRI having TLVs which do not have the correct lengths or syntax must be considered MALFORMED, and "treated-as-withdrawl".

* FSv2 NLRIs having TLVs which do not follow the above ordering rules described in section 4.1 MUST be considered as MALFORMED by a BGP FSv2 propagator, and treated "treated-as-withdrawl".

The above two rules prevent any ambiguity that arises from the multiple copies of the same NLRI from multiple BGP FSv2 propagators.

A BGP implementation SHOULD treat such malformed NLRIs as ‘session reset’ [RFC7606]

An implementation for a BGP speaker supporting both FSv1 and FSv2 MUST support the error handling for both FSv1 and FSv2.

4.2. Ordering of FS filters for BGP Peers which support FSv1 and FSv2

FSv2 allows the user to order flow specification rules and the actions associated with a rule. Each FSv2 rule has one or more match conditions and one or more actions associated with each rule.

FSv1 and FSv2 filters are sent as different AFI/SAFI pairs so FSv1 and FSv2 operate as ships-in-the-night. Some BGP peers in an AS may support both FSv1 and FSv2. Other BGP peers may support FSv1 or FSv2. Some BGP will not support FSv1 or FSV2. A coherent flow specification technology must have consistent best practices for ordering the FSv1 and FSv2 filter rules.

One simple rule captures the best practice: Order the FSv1 filters after the FSv2 filter by placing the FSv1 filters after the FSv2 filters.

To operationally make this work, all flow specification filters should be included the same data base with the FSv1 filters being assigned a user- defined order beyond the normal size of FSv2 user-ordered values. A few examples, may help to illustrate this best practice.

Example 1: User ordered numbering - Suppose you might have 1,000 rules for the FSv2 filters. Assign all the FSv1 user defined rules to 1,001 (or better yet 2,000). The FSv1 rules will be ordered by the components and component values.

Example 2: Storage of actions - All FSv1 actions are defined ordered actions in FSv2. Translate your FSv1 actions into FSv2 ordered actions for storing in a common FSv1-FSv2 flow specification data base.

5. Scalability and Aspirations for FSv2

Operational issues drive the deployment of BGP flow specification as a quick and scalable way to distribute filters. The early operations accepted the fact validation of the distribution of filter needed to be done outside of the BGP distribution mechanism. Other mechanisms (NETCONF/RESTCONF or PCEP) have reply-request protocols.

These features within BGP have not changed. BGP still does not have an action-reply feature.

NETCONF/RESTCONF latest enhancements provide action/response features which scale. The combination of a quick distribution of filters via BGP and a long-term action in NETCONF/RESTCONF that ask for reporting of the installation of FSv2 filters may provide the best scalability.

The combination of NETCONF/RESTCONF network management protocols and BGP focuses each protocol on the strengths of scalability.

FSv2 will be deployed in webs of BGP peers which have some BGP peers passing FSv1, some BGP peers passing FSv2, some BGP peers passing FSv1 and FSv2, and some BGP peers not passing any routes.

The TLV encoding and deterministic behaviors of FSv2 will not deprecate the need for careful design of the distribution of flow specification filters in this mixed environment. The needs of networks for flow specification are different depending on the network topology and the deployment technology for BGP peers sending flow specification.

Suppose we have a centralized RR connected to DDoS processing sending out flow specification to a second tier of RR who distribute the information to targeted nodes. This type of distribution has one set of needs for FSv2 and the transition from FSv1 to FSv2.

Suppose we have Data Center with a 3-tier backbone trying to distribute DDoS or other filters from the spine to combinational nodes, to the leaf BGP nodes. The BGP peers may use RR or normal BGP distribution. This deployment has another set of needs for FSv2 and the transition from FSv1 to FSV2.

Suppose we have a corporate network with a few AS sending DDoS filters using basic BGP from a variety of sites. Perhaps the corporate network will be satisfied with FSv1 for a long time.

These examples are given to indicate that BGP FSv2, like so many BGP protocols, needs to be carefully tuned to aid the mitigation services within the network. This protocol suite starts the migration toward better tools using FSv2, but it does not end it. With FSv2 TLVs and deterministic actions, new operational mechanisms can start to be understood and utilized.

This FSv2 specification is merely the start of a revolution of work not the end.

6. Optional Security Additions

This section discusses the optional BGP Security additions for BGP-FS v2 relating ROA [RFC9582].

6.1. BGP FSv2 with ROA

BGP FSv2 can utilize ROAs in the validation. If BGP FSv2 is used with BGPSEC and ROA, the first thing is to validate the route within BGPSEC and second to utilize BGP ROA to validate the route origin.

The BGP-FS peers using both ROA and BGP-FS validation determine that a BGP Flow specification is valid if and only if one of the following cases:

* If the BGP Flow Specification NLRI has a IPv4 or IPv6 address in destination address match filter and the following is true:
  - A BGP ROA has been received to validate the originator, and
  - The route is the best-match unicast route for the destination prefix embedded in the match filter; or
* If a BGP ROA has not been received that matches the IPv4 or IPv6 destination address in the destination filter, the match filter must abide by the [RFC8955] and [RFC8956] validation rules as follows:
  - The originator match of the flow specification matches the originator of the best-match unicast route for the destination prefix filter embedded in the flow specification", and
  - No more specific unicast routes exist when compared with the flow destination prefix that have been received from a different neighboring AS than the best-match unicast route, which has been determined in step A.

The best match is defined to be the longest-match NLRI with the highest preference.

7. IANA Considerations

This section complies with [RFC7153].

7.1. Flow Specification V2 SAFIs

IANA is requested to assign two SAFI Values in the registry at https://www.iana.org/assignments/safi-namespace from the Standard Action Range as follows:

Table 7-1 SAFIs

Value	Description	Reference
TBD1	BGP FSv2	[this document]
TBD2	BGP FSv2 VPN	[this document]

7.2. BGP Capability Code

IANA is requested to assign a Capability Code from the registry at https://www.iana.org/assignments/capability-codes/ from the IETF Review range as follows:

Table 7-2 - Capability Code

Value	Description	Reference	Controller
TBD3	Flow Specification V2	[this document]	IETF

7.3. Generic Transitive Extended Community

IANA is requested to assign a type value from the "Generic Transitive Extended Community Sub-Types" registry at https://www.iana.org/assignments/bgp-extended-communities/bgp-extended-communities.xhtml

Table 7-3 - Generic Transitive Extended Community

Value	Description	Reference	Controller
TBD4	FSv2 Action Chain Ordering	[this document]	IETF

The requested value is "0x01".

7.4. FSv2 IP Filters Component Types

IANA is requested to create a new "BGP FSv2 Component Types" registry and indicate [this draft] as a reference. The following assignments in the FSv2 IP Filters Component Types Registry shold be made.

Table 7-5 - Flow Specification

Registry Name: BGP FSv2 Component Types
Reference: [this document]
Registration Procedures: 0x01-0x3FFF Standards Action.

Value	Description	Reference
1	Destination filter	[RFC8955][RFC8956][this document]
2	Source Prefix	[RFC8955][RFC8956][this document]
3	IP Protocol	[RFC8955][RFC8956][this document]
4	Port	[RFC8955][RFC8956][this document]
5	Destination Port	[RFC8955][RFC8956][this document]
6	Source Port	[RFC8955][RFC8956][this document]
7	ICMP Type [v4 or v6]	[RFC8955][RFC8956][this document]
8	ICMP Code [v4 or v6]	[RFC8955][RFC8956][this document]
9	TCP Flags [v4]	[RFC8955][RFC8956][this document]
10	Packet Length	[RFC8955][RFC8956][this document]
11	DSCP marking	[RFC8955][RFC8956][this document]
12	Fragment	[RFC8955][RFC8956][this document]
13	Flow Label	[RFC8956][this document]

7.5. FSV2 NLRI TLV Types

IANA is requested to create the a new registries on a new "Flow Specification v2 TLV Types" web page.

This set of option assumes Type 2 ordering by packet frame rather than ascending order.

Table 7-6 FSv2 TLV types - Option 2
Registry Name: BGP FSv2 TLV types
Reference: [this document]
Registration Procedures: 0x01-0x3FFF Standards Action.

Type	Description	Reference
0x00	Reserved	[this document]
0x01-0x31	Unassigned	[this document]
0x32 (50)	L2 Traffic Rules	[this document]
0x33-0x63	Unassigned	[this document]
0x64 (100)	MPLS traffic rules	[this document]
0x65 (101)	Unassigned	[this document]
-0x95 (149)		
0x96 (150)	SFC Traffic rules	[this document]
0x97 (151)	Unassigned	[this document]
- 0xC7		
0xC8 (200)	Tunnel Traffic rules	[this document]
0xC9-0x99	Unassigned	[this document]
0x100 (256)	IP traffic rules	[this document]
0x101-0x117	Unassigned	
0x118 (280)	Extended IP Rules	[this document]
0x119-0x6000	Unassigned	[this document]
0x6000-0x7FFF	Vendor specific	[this document]
0x8000-0xFFFF	Reserved	[this document]

8. Security Considerations

The use of ROA improves on [RFC8955] by checking to see of the route origination. This check can improve the validation sequence for a multiple-AS environment.

>The use of BGPSEC [RFC8205] to secure the packet can increase security of BGP flow specification information sent in the packet.

The use of the reduced validation within an AS [RFC9117] can provide adequate validation for distribution of flow specification within a single autonomous system for prevention of DDoS.

Distribution of flow filters may provide insight into traffic being sent within an AS, but this information should be composite information that does not reveal the traffic patterns of individuals.

9. References

9.1. Normative References

[I-D.ietf-idr-flowspec-interfaceset]
Litkowski, S., Simpson, A., Patel, K., and J. Haas, "Applying BGP flowspec rules on a specific interface-set", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-interfaceset-06, 2 September 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-interfaceset-06>.

[I-D.ietf-idr-flowspec-l2vpn]
Hao, W., 3rd, D. E. E., Litkowski, S., and S. Zhuang, "BGP Dissemination of L2 Flow Specification Rules", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-l2vpn-27, 16 March 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-l2vpn-27>.

[I-D.ietf-idr-flowspec-nvo3]
Eastlake, D. E., Weiguo, H., Zhuang, S., Li, Z., and R. Gu, "BGP Dissemination of Flow Specification Rules for Tunneled Traffic", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-nvo3-23, 5 December 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-nvo3-23>.

[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>.

[I-D.ietf-idr-flowspec-redirect-ip]
Haas, J., Henderickx, W., and A. Simpson, "BGP Flow-Spec Redirect-to-IP Action", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-redirect-ip-06, 2 March 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-redirect-ip-06>.

[I-D.ietf-idr-flowspec-srv6]
Li, Z., Chen, H., Loibl, C., Mishra, G. S., Fan, Y., Zhu, Y., Liu, L., Liu, X., and S. Zhuang, "BGP Flow Specification for SRv6", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-srv6-08, 24 November 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-srv6-08>.

[I-D.ietf-idr-wide-bgp-communities]
Raszuk, R., Haas, J., Lange, A., Decraene, B., Amante, S., and P. Jakma, "BGP Community Container Attribute", Work in Progress, Internet-Draft, draft-ietf-idr-wide-bgp-communities-12, 17 March 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-wide-bgp-communities-12>.

[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>.

[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>.

[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>.

[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>.

[RFC5065] Traina, P., McPherson, D., and J. Scudder, "Autonomous System Confederations for BGP", RFC 5065, DOI 10.17487/RFC5065, August 2007, <https://www.rfc-editor.org/info/rfc5065>.

[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>.

[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>.

[RFC9015] Farrel, A., Drake, J., Rosen, E., Uttaro, J., and L. Jalil, "BGP Control Plane for the Network Service Header in Service Function Chaining", RFC 9015, DOI 10.17487/RFC9015, June 2021, <https://www.rfc-editor.org/info/rfc9015>.

[RFC9117] Uttaro, J., Alcaide, J., Filsfils, C., Smith, D., and P. Mohapatra, "Revised Validation Procedure for BGP Flow Specifications", RFC 9117, DOI 10.17487/RFC9117, August 2021, <https://www.rfc-editor.org/info/rfc9117>.

[RFC9184] Loibl, C., "BGP Extended Community Registries Update", RFC 9184, DOI 10.17487/RFC9184, January 2022, <https://www.rfc-editor.org/info/rfc9184>.

[RFC9582] Snijders, J., Maddison, B., Lepinski, M., Kong, D., and S. Kent, "A Profile for Route Origin Authorizations (ROAs)", RFC 9582, DOI 10.17487/RFC9582, May 2024, <https://www.rfc-editor.org/info/rfc9582>.

9.2. Informative References

[I-D.hares-idr-fsv2-more-ip-actions] Hares, S., "BGP Flow Specification Version 2 - More IP Actions", Work in Progress, Internet-Draft, draft-hares-idr-fsv2-more-ip-actions-03, 17 October 2024, <https://datatracker.ietf.org/doc/html/draft-hares-idr-fsv2-more-ip-actions-03>.

[I-D.hares-idr-fsv2-more-ip-filters]
Hares, S., "BGP Flow Specification Version 2 - More IP Filters", Work in Progress, Internet-Draft, draft-hares-idr-fsv2-more-ip-filters-04, 15 November 2024, <https://datatracker.ietf.org/doc/html/draft-hares-idr-fsv2-more-ip-filters-04>.

[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>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <https://www.rfc-editor.org/info/rfc8040>.

[RFC8205] Lepinski, M., Ed. and K. Sriram, Ed., "BGPsec Protocol Specification", RFC 8205, DOI 10.17487/RFC8205, September 2017, <https://www.rfc-editor.org/info/rfc8205>.

[RFC8206] George, W. and S. Murphy, "BGPsec Considerations for Autonomous System (AS) Migration", RFC 8206, DOI 10.17487/RFC8206, September 2017, <https://www.rfc-editor.org/info/rfc8206>.

Authors' Addresses

Susan Hares
Hickory Hill Consulting
7453 Hickory Hill
Saline, MI 48176
United States of America
Phone: +1-734-604-0332
Email: shares@ndzh.com

Donald Eastlake
Independent
2386 Panoramic Circle
Apopka, FL 32703
United States of America

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com

Jie Dong
Huawei Technologies
No. 156 Beiqing Road
Beijing
China
Email: jie.dong@huawei.com

Chaitanya Yadlapalli
ATT
United States of America
Email: cy098d@att.com

Sven Maduschke
Verizon
Germany
Email: sven.maduschke@de.verizon.com

Jeff Haas
Juniper
United States of America
Email: jhaas@juniper.net