IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 6, 2016

S. Hares
Huawei

March 5, 2016

An Information Model for Basic Network Policy and Filter Rules
draft-hares-idr-flowspec-combo-01.txt

Abstract

BGP flow specification (RFC5575) describes the distribution policy that contains filters and actions that apply when packets are received on a router with the flow specification function turned on. The popularity of these flow specification filters in deployment for DoS and SDN/NFV has led to the requirement for more BGP flow specification match filters in the NLRI and more BGP flow specification actions. Two solutions exist for adding new filters: 1) expanding the BGP Flow Specification version 1 (NLRI match filters and extended communities actions) to included limited number of filters and actions, and 2) creating a BGP Flow Specification version 2 that allows for ordering filters and actions (using new NLRI and wide-communities for actions). The two solutions can exist in parallel.

This document contains an overview existing proposals for expansion of BGP flow specification policy, proposals for BGP Flow Specification v1 and a new BGP Flow specification version 2 that supports order of filters and actions plus allowing more actions. This document also provides rules for the interaction of IDR Flow Specification policy (session ephemeral policy) with policy found in I2RS (reboot ephemeral policy), and policy found in ACLs and Policy routing (configuration policy). This document does not contain the individual definitions of policy rule conditions or actions.

Copyright Notice

Table of Contents

1. Introduction

BGP flow specification (RFC5575) describes the distribution of filters and actions that apply when packets are received on a router with the flow specification function turned on. If one considers the reception of the packet as an event, then BGP flow specification describes a set of minimalistic Event-MatchCondition-Action (ECA) policies. The initial set of policy (RFC5575 and RFC7674) for this policy includes 12 types of match filters encoded in the NLRI for two types of SAFIs (IP-only SAFI, 133; VPN SAFI, 134) for IPv4. The popularity of these flow specification filters in deployment for DoS and SDN/NFV has led to the requirement for more BGP flow specification match filters in the NLRI and more BGP flow specification actions.

Two solutions exist for adding new filters: 1) expanding the BGP Flow Specification (NLRI match filters and extended communities actions) for a limited number of filters and actions, and 2) creating a BGP Flow Specification version 2 that allows for ordering filters and actions (using new NLRI and wide-communities [I-D.ietf-idr-wide-bgp-communities] for actions). The two solutions

can exist in parallel. This document contains an overview of both solutions, rules for combining new flow specification policies which support IPv6, L2, nvo03 and MPLS match filters and new actions, and suggestions on how to expand yang modules to monitor both types. This document also provides rules for the interaction of IDR Flow Specification policy (session ephemeral policy) with policy found in I2RS (reboot ephemeral policy), and policy found in ACLs and Policy routing (configuration policy). This document does not contain the individual definitions of policies whcih are contained in the other specifications.

Section 1 of this draft contains an introduction to BGP flow specification [RFC5575] and drafts expanding the RFC5575 state. Section 2 contains the definitions related to this draft. Section 3 provides an overview of existing and proposed flow specification policy rules decribed in terms of packet event, packet match conditions, and actions (packet forwarding or packet match). The flow specification policies reviewed include policy in RFCs ([RFC5575], [RFC7674]), IDR WG documents ([I-D.ietf-idr-flow-spec-v6], [I-D.ietf-idr-flowspec-l2vpn]), and the following proposed IDR WG documents

o [I-D.eddy-idr-flowspec-packet-rate] (traffic limiting by packet rate),

o [I-D.eddy-idr-flowspec-exp] (Extensions for BGP security and others),

o [I-D.hao-idr-flowspec-nvo3] (flow specification for inner/outer nv03 forwarding),

o [I-D.hao-idr-flowspec-redirect-tunnel] (redirect to tunnel),

o [I-D.liang-idr-bgp-flowspec-label] MPLS label related filters and actions,

o [I-D.liang-idr-bgp-flowspec-time] Filters by time,

o [I-D.litkowski-idr-flowspec-interfaceset]Filters applied by order for Interface group, and

o [I-D.vandevelde-idr-flowspec-path-redirect]Filters applied to packet identifier,

Section 4 describes a proposal for an enhancement of BGP Flow specification security for both proosal. This security enhancement suggests using BGP ROA and allows the addition of BGP security to

validate the AS Path or AS Extended Communities and AS Wide Communities.

Section 5 describes the minimal subset solution with:

- o summary of NLRI and extended community formats (xection 5.1)

- o security addition of ROA (section 5.2),

- o match filter list and precedence of match filters (section 5.3),

- o action list and precedence of actions(section 5.4),

- o conflict with other Packet-reception Event-MatchCondition-Action (ECA) policy (I2RS Filter-Based RIB and Policy-Based Routing (n-tuple forwarding)) (section 5.9),

- o pros-cons of this approach (section 5.10)

Section 6 contains the BGP Flow specification with the sub-sections as section 4 except that section one summarizes the new NRLI with ordering of filters, and wide community atoms.

Section 7 proposes changes to the proposed Flow Specification Yang Module ([I-D.wu-idr-flowspec-yang-cfg]. yang modules in order to provide common monitoring of BGP Flow Specification version 1 and version 2. The changes suggest include changes to:

- o local configuration of BGP Flow Specification to be distributed to remote peers,

- o storage of bgp policy received from remote BGP peers [operational state],

- o statistics on use of locally configured BGP Flow Specification and remotely configured BGP Flow specification [operational state].

In addition, this section suggests ways to store BGP Flow Specification that will aid in comparing the BGP Flow Specification with other packet-reception ECA policy.

Section 9 discusses the security considerations for all the BGP Flow Specifications.

1.1. Overview of RFC5575

[RFC5575] describes the dissemination of flow specification rules via groups BGP Multi-Protocol NLRIs and BGP communities. A flow specification operates on packets received in a router when the flow specification feature is configured. The flow specification specifies match conditions for filters for packets received by a router and actions to do based on a match of those filters. If one considers the reception of a packet as an event, then a BGP flow specifications can be considered a set of minimalistic Event-Match Condition-Action policies (ECA policies). This set is minimalistic because there is only one event - the reception of a packet. BGP Flow specifications are BGP policy passed between peers.

The BGP flow specification policy is specified in filters contained in the MP-BGP NLRIs and actions contained within BGP Extended communities. The BGP peer propagates the flow-specifications between domains in order to automate inter-domain coordination of traffic filtering. Two applications that are using this are: distributed denial of service attack suppression and traffic filtering in BGP/ MPLS VPN service. BGP. BGP flow specifications use SAFI 133 non-VPN flow specifications, and SAFI 134 for BGP VPN flow specificatinos.

BGP Flow specification are validated based on:

a) originator of flow specification matching the originator of the best-match unicast route for the destination prefix embedded in the flow specification, and

b) no more specific unicast routes, when compared with flow destination prefix, that have been received from differenting neighboring AS than the best-match unicast route

Originator is specified by BGP originator path attribute or transport address of the BGP peer sending the BGP Flow specification. To support BGP flow specification, implementations are required to enforce the neighbor AS in the AS_PATH attribute is in the left-most position of AS_PATH.

```
        +-----------------------------+
        | Flow Specification (FS)     |
        |  Policy                     |
        +-----------------------------+
            ^                  ^
            |                  |
            |                  |
   +--------^-------+   +-------^-------+
   |   FS Rule      |   |   FS Rule     |
   +----------------+   +---------------+
                          :          :
                          :          :
                    ......:          :.....
                    :                     :
          +---------V---------+      +----V-------------+
          |  Rule Condition   |      |   Rule Action    |
          |  in BGP NLRIs     |      | in BGP extended  |
          | SAFI 133, 134     |      | Communities      |
          +-------------------+      +------------------+
              :     :    :                 :     :    :
         .....:     .    :.....       .....:     .    :.....
         :          :         :       :          :         :
    +----V---+  +---V----+ +--V---+ +-V------++--V-----++--V---+
    |  Match |  | match  | |match | | Action || action ||action|
    |Operator|  |Variable| |Value | |Operator||Variable|| Value|
    |*1      |  |        | |      | |(type-) ||        ||      |
    +--------+  +--------+ +------+ +--------++--------++------+
```

*1 match operator for Types 3-12. Match operator supports pairs of matching operators.

Figure 1: BGP Flow Specification Policy

Match operators includes a sequence of match operations each with the form [op, value] where match can match values greater, lessthan, or equal to teh value. The sequence of match operators can be combined as logical AND or ORs.

1.2. Flow Specifications: Ephemeral or not?

BGP Flow specification does not indicate what happens to the flow specifications if a BGP peering session closes. [RFC5575] specifies a link to received "best-match" unicast routes, but does not provide any standard way of determining whether the flow specification sent by the BGP peer is kept after the BGP session closes. It is unclear whether BGP Flow specifications disappear when a BGP session closes (denoted as BGP session ephemeral), or disapppear when the BGP

module's hardware or software reboots (reboot ephemeral), or it is kept like configuration state that survives a reboot.

This document specifies that the default policy is that the BGP Flow Specification received from remote peers like other BGP peer state received from remote peers disappears when the BGP peer session closes. Local BGP Peer configuration is like all local configuration and persists while the BGP Peer is configured.

If an implementation decides to implement operator-applied policy that retains remotely received BGP Flow Specification policy after the BGP Peer closes, this action must be treated as if these BGP Flow Specification policy was locally configured. Therefore, these two actions are out of scope of this document.

1.3. Precedence between BGP Flow Specification and other packet-ECA policies

Why is this precedence bewteen BGP Flow Specification and other packet-ECA policies needed?

[RFC5575] states that Flow specification takes advantage of the "ACL" feature (section 1), but it does not state how BGP Flow specification interacts with ACL features. NETCONF [RFC6241] or RESTCONF [I-D.ietf-netconf-restconf] can be used to set ACL configuration state using the [I-D.ietf-netmod-acl-model] yang data module.

One of the proposals for a new BGP Flow specification action ([I-D.litkowski-idr-flowspec-interfaceset]) proposes an action which defines that a specific ordering of BGP flow-specifications and ACLs interaction for a set of interfaces for the drop/forward actions (see section 3 for details). This action proposals suggests a precedence between these two filter actions.

ACL is not the only packet-ECA policy used as an alternative to destination based routing. Two other n-tuple packet-reception ECA modules exist: n-tuple policy-based RIB/FIB (aka policy routing) and I2RS Filter-based RIB. The n-tuple policy based forwarding RIB/FIB configured on specific interfaces, and forward based on the match of an n-tuple filter that modifies, forwards, or drop n-tuples. If no match exists, this packet-reception ECA RIB forward this to a default RIB. A proposal for standardized yang model for this is in (draft-rtgwg-hares-rtgwg-fb-rib-00.txt).

The I2RS Filter-Based RIB (FB-RIB) also specifies another way to do flow filtering per packet/frame being received (n-tuple packet ECA policy) ([I-D.kini-i2rs-fb-rib-info-model], [I-D.hares-i2rs-fb-rib-data-model]) using a packet filter event-

match_condition-action policy [I-D.hares-i2rs-pkt-eca-data-model]. The I2RS protocol allows a I2RS Client to talk to an I2RS Agent within a routing device ([I-D.ietf-i2rs-architecture]) to set ephemermal policy which is module ephemeral and box ephemeral. The I2RS match_conditions examine frame/packet information (L1-L4, NV03, and SFC), and I2RS match_actions that modify packet/frame information. Figure 2 shows the structure of packet filtering ECA rules from [I-D.hares-i2rs-fb-rib-data-model] which used by I2RS Filter-Based RIB (FB-RIB). Note that these I2RS Filters have each rule has policy rule name, policy rule order number, and rule status.

Section 5 compares the filters and actions between BGP Flow Specification, I2RS Filter-Based RIB, Filter-RIB (aka Policy-Based Routing), and the ACL. The I2RS packet filter rules also allow the rule to be ordered and named. I2RS flow-based filters are ephemeral state [I-D.ietf-i2rs-ephemeral-state] are stored as ephemeral state which is lost upon a reboot.

```
    +-----------+     +------------+
    |Rule Group |     | Rule Group |
    +-----------+     +------------+
         ^                   ^
                 |                   |
         |                   |
+--------^-------+   +-------^-----------+
|      Rule      |   |     Rule          |
+----------------+   +-------------------+
                      :  :   :    :
    :.................:  :   :    :
    :          |.........:   :    :
 +--V--+    +--V--+          :    :
 | name|    |order| .........:    :.....
 +-----+    +-----+ :                  :
                    :                  :
     +--------------V-------+       +--V------------+
         | Rule Match condition |       | Rule Action   |
     +---------------------+       +---------------+
           :     :    :                 :     :    :
      .....:     .    :.....       .....:     .    :.....
      :          :         :       :          :         :
 +----V---+  +---V----+ +--V---+ +-V------++--V-----++--V---+
 |  Match |  | match  | |match | | Action || action ||action|
 |Operator|  |variable| |Value | |Operator||Variable|| Value|
 +--------+  +--------+ +------+ +--------++--------++------+
```

Figure 2: I2RS Filter-Based RIB Policy

1.4. BGP Flow Specification and logging

[RFC5575] specifies the Traffic Action Extended Community which specifies a Terminal (T) action flag and Sampling (S) flag. The sample flag indicates that "traffic sampling and logging" [is enabled] for a set of flow specifications in a BGP packet. the details of traffic sampling and logging are not specified in this standard. Logging and sampling provide valuable information to establish the impact of BGP Flow specification in order to automatic intra-AS DoS prevention or inter-AS automation of DOS or VPN traffic filters. [RFC5575] was written before the advent of yang modules that specify operational state [I-D.ietf-netmod-opstate-reqs]. [I-D.wu-idr-flowspec-yang-cfg] proposes a BGP Flow Specification Yang Data model with BGP Flow Specification configuration, operational state for BGP Flow specifications received from peers (BGP Session Ephemeral state), and statistics on the use of filters, actions, and dropped packets. Section 7 describes how the logging and notifications for BGP Flow specifications can be added to this yang module.

1.5. BGP Flow Specification and BGPSEC

[RFC5575] does not require BGP Flow specifications to be passed BGPSEC [I-D.ietf-sidr-bgpsec-protocol]. [RFC5575] states "as long as traffic filtering rules are restricted to match the corresponding unicast routing paths for relevant prefixes, the security characteristics of this protocol are equivalent to existing security properties of BGP unicast properties", and "where this is not the case, this would open the door to further denial of service attack" (section 10). [I-D.eddy-idr-flowspec-exp] suggests passing BGP Flow Specification in BGPSEC. Section 10 summarizes the security issues with the current [RFC5575] and the enhancements described in this draft, and discusses the proposed fixes that that [I-D.eddy-idr-flowspec-exp] provides.

2. Definitions

2.1. Definitions and Acronyms

NETCONF: The Network Configuration Protocol [RFC6241].

RESTconf - http programmatic protocol to access yang modules [I-D.ietf-netconf-restconf]

BGPSEC - secure BGP [I-D.ietf-sidr-bgpsec-protocol].

I2RS - Interface to Routing System [I-D.ietf-i2rs-architecture].

ephemeral - state which does not survive a particular event.

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

Reboot ephemeral state - state which does not survive the reboot of a software module, or a hardware reboot.

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

2.2. 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 [RFC2119].

3. BGP Flow Specification Policy - Original and Expansions

3.1. Packet Reception Event

The reception of a packet is the event that causes the BGP policy to enact. By default the BGP Flow specification applies to all interfaces. This can be restricted by a BGP Flow Specification Action or policy local to a node running the BGP peer session.

The definition of a packet is not limited to a IP packet (IPv4 or IPv6) but also includes mpls packets, L2 frames (802.1Q), encapsulated packets (NVGRE or VXLAN or any other NV03 encapsulation).

The same definition of the event is utilized by the I2RS Filter-based RIBs ([I-D.kini-i2rs-fb-rib-info-model] and [I-D.hares-i2rs-fb-rib-data-model] and the Filter-Based RIBs (draft-hares-rtgwg-fb-rib-data-model), and ACL filters [I-D.ietf-netmod-acl-model].

These packet events are the standardized packet events. Additional packet events for vendors may augment these standards events.

3.2. BGP Flow Specification Match Filters

[RFC5575] defines match conditions for IPv4 to be carried with the NLRI format for 12 types of packet match events (see figure 3), and that all filters specified must be combined by a "AND". The proposed expansions to this filter list utilizing the Flow Specification NLRI are listed in figure 4. [I-D.li-idr-flowspec-rpd] proposed a BGP

Attribute which contains additional flow specification filters, and actions. Figure 5 contains the match filters from this draft.

The proposals to expand flow specification beyond [RFC5575] filter specifications include:

Matches for the inner-outer header for encapsulated traffic for being specified for the NV03 networks (MF-1, MF-2, MF-3) in [I-D.hao-idr-flowspec-nvo3],

extended match filters carried in BGP attribute which includes time (MF-5) for enacting flow-specification filter rules ([I-D.li-idr-flowspec-rpd], [I-D.liang-idr-bgp-flowspec-time]).

One filter that seems obvious is the filter for the MPLS labels. However, no proposal includes this Match filter for MPLS.

The precedence order for the match filter rules was specified in [RFC5575] and expanded in [I-D.ietf-idr-flowspec-l2vpn]. The combined precedence is shown in figure 4.

Table 1: IDR WG BGP Flow Specification Match Filter

type#	Type Name	Match	Reference
1	Destination Prefix	IPv4 Prefix	RFC5575
		IPv6 Prefix	ietf-idr-flow-spec-v6
2	Source Prefix	IPv4 Prefix	RFC5575
		IPv6 Prefix	ietf-idr-flow-spec-v6
3	IP protocol	IPv4 Protocol	RFC5575
		number	
3	Next Header	IPv6 protocol	ietf-idr-flow-spec-v6
4	Port (source or	Port number	RFC5575
	destination port)		RFC5575
5	Source port	Port number	RFC5575
6	Destination port	Port number	RFC5575
7	ICMP type	ICMP type	RFC5575
8	ICMP code	ICMP code	RFC5575
9	TCP Flags	1 or 2 byte	RFC5575
		bitmask for	RFC5575
		TCP flags	
10	Packet length	# of bytes	RFC5575
	(for IP packet)		
11	DSCP	IPv4 DSCP	RFC5575
		(6 bit mask)	RFC5575
11	Traffic class	IPv6 traffic	ietf-idr-flow-spec-v6
		(8 bit mask)	
12	IPv4 Fragment	4 bit mask	RFC5575
13	IPv6 Flow	20 bit flow	ietf-idr-flow-spec-v6
14	Ethernet type	2 bytes	ietf-idr-flowspec-l2vpn
15	Source MAC	MAC address	ietf-idr-flowspec-l2vpn
16	Destination MAC	MAC Address	ietf-idr-flowspec-l2vpn
17	DSAP in LLC	1 octet	ietf-idr-flowspec-l2vpn
18	SSAP in LLC	1 octet	ietf-idr-flowspec-l2vpn
19	LLC Control field	1 octet	ietf-idr-flowspec-l2vpn
20	SNAP	5 octets	ietf-idr-flowspec-l2vpn
21	VLAN ID	1 or 2 bytes	ietf-idr-flowspec-l2vpn
22	VLAN COS	3 bit COS	ietf-idr-flowspec-l2vpn
23	Inner VLAN ID	1 or 2 bytes	ietf-idr-flowspec-l2vpn
24	Inner VLAN COS	1 or 2 bytes	ietf-idr-flowspec-l2vpn

Figure 3

Table 2: Proposed BGP Flow Specification Match Condition Filters

type#	Type Name	Match	Reference
MF-1	Delimiter type (Encapsulation type VXLAN or NVGRE)	2 bytes	hao-idr-flowspec-nv03
MF-2	VNID (virtual network ID)	24 bit VN	hao-idr-flowspec-nv03
MF-3	Flow ID (NVGRE Flow ID)	8 bit flow ID	hoa-idr-flowspec-nv03
MF-4	MPLS LSP (label 20 bits, EXP (3 bits), S Bit TTL (8 bits)	TBD Label stack	not specified
MF-5	Interface (Group ID, intf id)	TBD	not specified

Figure 4

Table 3: Proposed BGP Flow Specifications Match in BGP Attribute

type#	Type Name	Match	Reference
MF-6	Time	??	liang-idr-bgp-flowspec -time

Figure 5

3.2.1. Current Precedence logic

Precedence logic for BGP Flow Specifications
(RFC5575, draft-idr-bgp-flowspec-l2vpn)

```
flow-rule-cmp (a,b)
{
  comp1 = next_component(a);
  comp2 = next_component(b);
  while (comp1 || comp2) {
   // component_type returns infinity on end of list
   if (component_type(comp1) < component_type(comp2)) {
    return A_HAS_PRECEDENCE;
    }

   if (component_type(comp1) > component_type(comp2)) {
    return B_HAS_PRECEDENCE;
   }

   // IP values)
   if (component_type(comp1) == IP_DESTINATION || IP_SOURCE) {
      common = MIN(prefix_length(comp1),prefix_length(comp2));
          cmp = prefix_compare (comp1,comp2,common);
          // not equal, lowest value has precedence
          // equal, longest match has precedence;
   } else if (component_type (comp1) == MAC_DESTINATION ||
              MAC_SOURCE) {
                common = MIN(MAC_address_length(comp1),
                             MAC_address_length(comp2));
                cmp = MAC_Address_compare(comp1,comp2,common);
                //not equal, lowest value has precedence
                //equal, longest match has precedence
        } else {
        common = MIN(component_length(comp1),
                             component_length(comp2));
            cmp = memcmp(data(comp1), data(comp2), common);
                //not equal, lowest value has precedence
                //equal, longest string has precedence
    }
  }
}
```

Figure 6

3.2.2. Why Current Match precedence Logic a problem

The current precedence logic requires the following:

o destination address (0/0 is fine for destination match,

- o components to go in numerical order,

- o and the matches to be an "AND of all component matches.

This does not allow matching MPLS before IP address, or MAC Addresses before IP addresses. This may make some n-tuple filter policies more difficult or even impossible to express in this fasion.

3.3. BGP Flow Specification Actions

[RFC5575] also defines four actions which would be carried in BGP extended communities: traffic rate (in bytes), traffic action, redirect to IPv4 VPAN, and traffic marking. Traffic action has two bits Terminal bit (T) and Sample (S) bit. If the Terminal Bit is set, the the node apply all filter rules based as defined by "AND" and precedence. If the terminal bit is clear, then the flow specification process is to stop. The Sample bit implies that the flow specification enables sampling and logging for this event.

Unfortunately, [RFC5575] was unclear about the "redirect to IP VPN action" and did not handle IPv6. [RFC7674] was written to clarify [RFC5575] by clearly specifying the 3 extended communities that "IPv4 VPN" needed to support AS 4 byte, and IPv4 address Routing Distinguishers (RDs). [I-D.ietf-idr-flow-spec-v6] was written to extend this work to IPv6 filters, and to include the IPv6 flow in the filter set as figure 5 shows.

Table 4: BGP Flow Specifications in RFC5575 and RFC7674

type#	Action name	action	Reference
0x8006	Traffic Rate (in bytes)	2 octet AS 4 octet float	RFC5575
0x8007	Traffic Action (S:Sample and log, T:last flowspec	6 octet bit mask:S,T bits	RFC5575
0x8008	Redirect (IP VPN) (RD: 2 octet AS, 4 octet value)	Route Target (6 octet)	RFC5575 and RFC7674
0x8108	Redirect (IP VPN) (RD: 4 octet IPv4 address, 2 byte value)	Route Target (6 octet)	RFC7674
0x8208	Redirect (IP VPN) (RC: 4 byte AS, 2 byte value)	Route Target	RFC7674

Figure 7

3.3.1. Proposals to extend these standardized actions

Proposals to extend the actions take upon a match include:

- o (FA1) [I-D.eddy-idr-flowspec-packet-rate] specifies a traffic rate limit by packets the number of packets forwarded,
- o (FA2)[I-D.li-idr-flowspec-rpd] specifies an "R" bit for traffic action that allows a BGP Attribute to pass additional BGP Flowspecification match filters and actions,
- o (FA3) [I-D.hao-idr-flowspec-redirect-tunnel] specifies a redirection to a tunnel specified in [I-D.rosen-idr-tunnel-encaps],
- o (FA4)[I-D.ietf-idr-flowspec-l2vpn] specifie push, pop, or swap VLANs before forwarding,
- o (FA5) [I-D.ietf-idr-flowspec-l2vpn] specifies the ability to replace TPIDs values with new values before forwarding,
- o (FA6) [I-D.liang-idr-bgp-flowspec-label] specifies push/pop/swap on MPLS labels before forwarding,

o (FA7)[I-D.litkowski-idr-flowspec-interfaceset] which specifies that ACL filters plus BGP flow specification filters will determine the acceptance/drop of inbound packet, and the forwarding/drop of outbound packets.

Figure 8 shows these flow specifications.

Table 5: Proposed Flow Specification Actions

type#	Action name	action	Reference
FA1	Traffic Rate (in packets)	2 octet AS 4 octet float	eddy-idr-flowspec-packet-rate
FA2	Extended Traffic Extension for R to take additional Flow specifications from BGP Flow spec Policy attribute	R bit P bit	li-idr-flowspec-rpd Alternate action procedures(this draft)
FA3	Redirect to tunnel (tunnel in BGP Attribute)	6 octets 1 bit flag (C=applies to copies only)	hao-idr-flowspec-redirect-to-tunnel
FA4	VLAN-action (push, pop, swap)	bitmask	idr-bgp-flowspec-l2vpn
FA5	TPID Action (NVGRE Flow ID)	6 octets	idr-bgp-flowspec-l2vpn
FA6	Label Action (push/pop/swap MPLS label uses Exp flag, TTL, Stack flag (S))	MPLS Tag, TTL(1 octet) S bit	liang-idr-bgp-flowspec-label-01
FA7	Alternate NLRI Validation (mask for support of RFC5755, ROA and bgpsec-protocol AS path) and L2MAC NRLI for IP Address	validation bit mask	eddy-idr-flowspec-exp (some functions)
FA8	for Interface set filter ACL + Flow specification rules	4 Byte AS 2 byte interface	litkowski-idr-flowspec-interfaceset

```
    |       |                    | group ID     |                      |
    +=======+====================+=============+======================+

    Note: FA8 is really a filer plus an action:
     FA8-filter: Restrict processing for filters to set of interfaces
     FA8-Action: Forward only if: ACL + Flow-Specification filters
                 suggest forwarding.
```

Figure 8

3.3.2. Why ordering is needed

One the probems with adding the actions is that precedence has not been set for the actions, and some actions can conflict. (see section

[RFC5575] indicates that the actions specified in the document represent only the "subset of filtering actions that can be interpreted across the network". As additional standardized actions occur, the non-standard action will need to have a precedence below the standardized actions.

To allow better security for Flow Specification NLRIs, the BGP validation of prefixes using the Route Origination (ROAs) technology ([RFC6483]) should be placed as the first action for a prefix. If the path needs to be validated The bgp-sec protocol [I-D.ietf-sidr-bgpsec-protocol] can be used to validate the AS path and actions. These validations must be first, and this is not allowed with the current actions.

One the probems with adding the actions is that precedence has not been set for the actions, and some actions may conflict. Table 6 suggests an order with the fewest conflicts, but even there proposal will need to be updated to handle these conflicts.

Table 6 - Action Precedence and Conflicts between Actions

order	Action	Possible Conflicting Actions
FA7	Alternate NLRI	none
1	Validation (mask for support of RFC5755, ROA and bgpsec-protocol AS path)	
2	Traffic Rate(0x8006)	Traffic rate in packets (FA1)

```
    |     | in bytes             |                                  |
    |     |                      | Default Conflict action:         |
    |     |                      | Allow traffic monitoring by bytes|
    |     |                      | and packets, but process byte    |
    |     |                      | rate limit checks first          |
    |     |                      |                                  |
    |  3  | Traffic Rate (FA1)   | traffic rate in bytes (0x8006)   |
    |     | in packets           |                                  |
    |     |                      | Default Conflict action: same    |
    |     |                      | as in Traffic Rate action        |
    |     |                      | conflict                         |
    |     |                      |                                  |
    |  4  | Traffic Action       | Extended Traffic action with     |
    |     |  (0x8007)            | "R-Policy" bit(FA2), "TN-P" bit, |
    |     |                      |  R-intf bit                      |
    |     |                      |                                  |
    |     |                      | Default conflict action: Process |
    |     |                      | Traffic Action, then Extended    |
    |     |                      | traffic action                   |
    |     |                      |                                  |
    |  5  | Extended Traffic     | Traffic Action (0x8007)          |
    |     | Action (FA2)         |"R" bit(FA2), "TN-P" bit (above)  |
    |     |                      | R-Intf bit                       |
    |     |                      |                                  |
    |     |                      | Default conflict action: Process |
    |     |                      | Traffic action, then extended    |
    |     |                      | traffic action                   |
    |     |                      |                                  |
    |  6  | Redirect to IP-VPN   | Redirect to IP Tunnel (FA3)      |
    |     | 0x8008: 2 byte AS RD| VLAN-action (FA4),               |
    |     | 0x8108: 4 byte IP RD| TPID-action (FA5)                |
    |     | 0x8208: 4 byte AS RD| Label-action (FA6)               |
    |     |                      | interface set (FA7)              |
    |     |                      |                                  |
    |     |                      | Default Conflict action:         |
    |     |                      | Process forward to IP-VPN first  |
    |     |                      | and ignore other conflicting     |
    |     |                      | actions unless TN-Mod bit set in |
    |     |                      | Extended action.
    |     |                      | If TN-Mod set then process the   |
    |     |                      | conflict actions which change    |
    |     |                      | the packet prior to forwarding   |
    |     |                      | the packet via tunnel to IP-VPN. |
    |     |                      |                                  |
    |     |                      | If I bit set, process interface  |
    |     |                      | restriction's narraowing of scope|
    |     |                      | to certain interfaces before     |
    |     |                      | processing other options, and    |
```

```
   |     |                      | process interface restrictions   |
   |     |                      | implied in outboudn direction    |
   |     |                      | before sending packet.           |
   |     |                      | outbound policy before any other |
   |     |                      | If "R" bit set use version 2 of  |
   |     |                      | BGP Flow Specification handling  |
   |     |                      |                                  |
   |  7  | Redirect to IP       | Redirect to IP VPN (0x8008,      |
   |     | Tunnel (FA3)         | 0x8108, 0x8208)                  |
   |     |                      | VLAN-action (FA4),               |
   |     |                      | TPID-action (FA5),               |
   |     |                      | Label action (FA6),              |
   |     |                      | interface set (FA7)              |
   |     |                      |                                  |
   |     |                      | Default Conflict actions:        |
   |     |                      | Refer to processing in redirect  |
   |     |                      | IP-VPN tunnel                    |
   |     |                      |                                  |
   |  8  | VLAN action (FM4)    | Redirect to IP-VPN (0x8008,      |
   |     |                      | 0x8108, 0x8208),                 |
   |     |                      | Redirect to tunnel (FA3),        |
   |     |                      | VLAN-action (FA4),               |
   |     |                      | TPID-action (FA5),               |
   |     |                      | Label action (FA6),              |
   |     |                      | interface set (FA7)              |
   |     |                      |                                  |
   |     |                      | Default Conflict actions:        |
   |     |                      | Refer to processing in redirect  |
   |     |                      | IP-VPN tunnel                    |
   |     |                      |                                  |
   |  9  | TPID action (FM5)    | Redirect to IP-VPN (0x8008,      |
   |     |                      | 0x8108, 0x8208),                 |
   |     |                      | Redirect to tunnel (FA3),        |
   |     |                      | VLAN-action (FA4),               |
   |     |                      | TPID-action (FA5),               |
   |     |                      | Label action (FA6),              |
   |     |                      | interface set (FA7)              |
   |     |                      |                                  |
   |     |                      | Default Conflict actions:        |
   |     |                      | Refer to processing in redirect  |
   |     |                      | IP-VPN tunnel                    |
   |     |                      |                                  |
   | 10  | Label Action (FM6)   | Redirect to IP-VPN (0x8008,      |
   |     |                      | 0x8108, 0x8208),                 |
   |     |                      | Redirect to tunnel (FA3),        |
   |     |                      | VLAN-action (FA4),               |
   |     |                      | TPID-action (FA5),               |
   |     |                      | Label action (FA6),              |
```

		interface set (FA7) Default Conflict actions: Refer to processing in redirect IP-VPN tunnel
11	interface Set (FM8a)	Redirect to IP-VPN (0x8008, 0x8108, 0x8208), Redirect to tunnel (FA3), VLAN-action (FA4), TPID-action (FA5), Label action (FA6), Default Conflict actions: Refer to processing in redirect IP-VPN tunnel
12	Filter precedence (FM8b) [proposed]	reorder default filter precedence 0 = BGP Flow-Spec only 1 = ACL + BGP Flow-Spec 2 = I2RS FB-RIB + BGP FS 3 = ACL + I2RS FB-FIB + BGP FS 4 = Config FB-RIB + BGP FS 5 = ACL + config FB-RIB + BGP FS 6 = Config FB-RIB + I2RS FB-RIB + BGP FS 7 = ACL + config FB-FIB + I2RS
13-63		Reserved for other standards actions
65+	FCFS actions	FCFS Actions

Figure 9

Conflict process may have an ordering of the conflict processes or parallel processes. Due to this conflict processing also needs to have common diagrams or a language for precedence that is common across all rules. An example of a conflict diagram is below. Conflict 1 and Conflict 2 are parallel conflict resolutions that are run prior to conflict 3.

```
     action               precedence 1        precedence 2
    +----------+       +-----------+
    | action 1 |-------|conflict 1 |----|
    |          |       +-----------+    |   +----------+
    |          |                        |---|conflict 3|
    |          |       +-----------+    |   +----------+
    |          |-------|conflict 2 |----|
    +----------+       +-----------+

     precedence of conflicts for action 1 {}
      precedence(1) = conflict 1 | conflict 2;
      precedence(2) = conflict 3;
      If precedence (1) found; continue
      if precedence (3) found; exit;
     }
```

Figure 10

4. Proposal to Expand BGP Flow Specification Security

[RFC5575] does not require BGP ROA [RFC6483] as the BGP ROA was not standardized until after [RFC5575]. [RFC5575] states "as long as traffic filtering rules are restricted to match the corresponding unicast routing paths for relevant prefixes, the security characteristics of this protocol are equivalent to existing security properties of BGP unicast properties", and "where this is not the case, this would open the door to further denial of service attack" (section 10).

[RFC5575] requires an extension of the BGP route selection procedures [RFC4271] in section 9.1.2 in order to validate the BGP flow specification NLRI. The BGP Flow Specification NLRI is valid if and only if:

- o "the originator of the flow specification matches the orginator of the the best-match unicast route for the destination prefix embedded in the flow specification",

- o "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".

This set of validation requirements also require that BGP implementations are required to enforce the AS_PATH attribute having the neighbor AS in the left-most position.

4.1. Validation for NLRI with L2VPN validation

These validation steps required a unicast IPv4 or IPv6 route be transmitted with L2VPN ([I-D.ietf-idr-flowspec-l2vpn]) and the NV03 flow specifications [I-D.hao-idr-flowspec-nvo3] to validate the path. These specifications do not provide additional details on any additional validation needed for the L2VPN or NV03 Case.

4.2. Using ROA to validate BGP Flow Specification

Since [RFC5575] BGP Route Origin validation [RFC6482] has been standardized, and the BGPSEC protocol [I-D.ietf-sidr-bgpsec-protocol] has been developed. This document proposes that an action be created in both the proposals that has precedence over all other actions.

[I-D.eddy-idr-flowspec-exp] specifies cryptographic enhancements that include:

o creating a BGP identifier (in BGP attribute or in BGPSEC signature),

o Expanding BGPSEC coverage for Route Orgination Authorization (ROA) to cover the orignator of the BGP Flow specification for the BGP Flow specification SAFIs.

o Covering the BGP Extended Communities with BGP signature.

While this work is interesting, the authors of [I-D.eddy-idr-flowspec-exp] consider it research into the use of BGP security. Therefore, this proposal suggest this addition be covered as an expansion to the ROA process. As this solitifies the ROA-action should be updated to include this functionality.

4.3. Using BGPSec to validate AS Path

The use of bgpsec protocol to validate the AS Path is orthongonal to the validation of the prefix to origin AS. Therefore, local configuration can determine if the bgpsec protocol is supported and required to validate the AS Path checked for the set of peers using BGP Flow Specification. If bgpsec is configured to be used, the BGP FLOW Specification SHOULD use the secured AS Path for its validation checks.

5. Minimal BGP-FS Additions (Option 1)

This section on minimal subset solution has:

summary of NLRI and extended community formats (xection 5.1)

security addition of ROA (section 5.2),

match filter list and precedence of match filters (section 5.3),

action list and precedence of actions(section 5.4),

conflict with other Packet-reception Event-MatchCondition-Action (ECA) policy (I2RS Filter-Based RIB and Policy-Based Routing (n-tuple forwarding)) (section 5.9),

pros-cons of this approach (section 5.10)

It is important to note that BGP Flow Specification is not the only packet reception ECA policy in a system. BGP Flow specification is session ephemeral state which is not guaranteed to persist when the BGP peer session closes. I2RS Filter-Based RIB is reboot ephemeral state which will not persist when the routing entity reboots. Policy RIB (aka Filter Forwarding RIB) and ACLs are configuration state which can persist over the reboot of a system. In many systems, operator-applied policy may set the priority between these systems. In order to provide interoperability between BGP Flow Specifcastion and current IETF management systems using yang-models accessed by netconf, restconf, and I2RS protocols, it important to define the default precedence between these different packet reception ECA policies. Section 5.9 provides the details on this proposals.

5.1. Summary of Existing Flow Specification Formats

The existing BGP Flow Specification is contained with the the BGP Flow Specification NLRI encoded using MP_REACH_NLRI and the MP_UNREACH_NLRI as defined in [RFC4760]. If the application does not require the next-hop field, it will be encoded as 0 length. The BGP FLow Specification NLRI is encoded as shown in figure 11. [RFC5575] specifies SAFI 133 for "dissemination of IPv4 flow specification", and SAFI 134 for "dissemination of VPNv4 Flow Specification". [I-D.ietf-idr-flow-spec-v6] expands the use of these SAFI to the IPv6 AFI. [I-D.ietf-idr-flowspec-l2vpn] expands this use to L2VPN for the VPLS [RFC4761], EVPN and LDP-Based VPLS [RFC4762] with BGP auto-discovery [RFC6074].

```
+-------------------------+
| length (0xnn or 0xfn nn)| (1 or 2 octets depending on encoding)
+-------------------------+
| NLRI Value (variable)   |
+-------------------------+

SAFI   AFIs
133    IPv4 (AFI=1),
       IPv6 (AFI=2)
134    IPv4 VPNs (AFI=1),
       IPv6 VPNs(AFI=2),
       L2VPN (AFI=25)
```

Figure 11

The actions for the BGP Flow Specification are carried in 6 bytes of the BGP Extended Community.

5.2. New Validation Rules for BGP Flow Specification: Precedence with ROA

This precedence within BGP Session Ephemeral state depends on the preference associated with valid BGP Session flow specification NLRI received within a BGP State. Since [RFC5575] was published, additional mechanisms to validate originating prefixes with an AS with Prefix Orgin Validation (ROA), and the BGPSEC Secure Path have been standardized. The precedence of these mechanisms should be from BGP Security to ROA to [RFC5575]. The BGP peers determine that a BGP Flow specification is valid if and only if one of the following cases:

- o 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

- o 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 [RFC5575] validation rules of:

  * 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.

5.3. Match Condition Filters with Precedence Ordering

Match conditions depends on an "AND" of all rules within a Flow Specification policy. A Flow specification policy is defined by a sequence of BGP Flow specification NLRIs with filter-match rules. The sequence of Flow Specification rules are terminate Traffic Action with a T-Bit flag set to zero.

Match condition processing occurs in the following overall precedence ordered from IP protocol to

1. IP Protocol (1-13),

2. NV03-matches (MF-1 to MF-3),

3. Other overlay matches (spring, SFC)

4. L2VPN matches (14-24),

5. MPLS matches (MF-4),

6. L2VPN matches (currently 14-24),

7. interfaces matches (MF-5),

8. time matches (MF-6), and

9. Non-Standardized (First-Come-First Serve(FCFS)) match conditions (see [RFC5575] section 11)

Editorial note: This list is longer than many, and will be discussed on the IDR mail list.

Table 6 in figure 9 shows the filter by filter precedence order. All flow specification filters combine as an "AND" of all filters. A re-ordering of match filters is only possible in the the proposed version 2 of BGP Flow specification.

5.3.1. Table of Match Filters and Precedence

Table 8: Flow Specification Match Filter Precedence Order

type#	Type Name	Match	Reference
1	Destination Prefix	IPv4 Prefix	RFC5575
		IPv6 Prefix	ietf-idr-flow-spec-v6
2	Source Prefix	IPv4 Prefix	RFC5575
		IPv6 Prefix	ietf-idr-flow-spec-v6
3	IP protocol	IPv4 Protocol	RFC5575
		number	
3	Next Header	IPv6 protocol	ietf-idr-flow-spec-v6
4	Port (source or	Port number	RFC5575
	destination port)		RFC5575
5	Source port	Port number	RFC5575
6	Destination port	Port number	RFC5575
7	ICMP type	ICMP type	RFC5575
8	ICMP code	ICMP code	RFC5575
9	TCP Flags	1 or 2 byte	RFC5575
		bitmask for	RFC5575
		TCP flags	
10	Packet length	# of bytes	RFC5575
	(for IP packet)		
11	DSCP	IPv4 DSCP	RFC5575
		(6 bit mask)	RFC5575
11	Traffic class	IPv6 traffic	ietf-idr-flow-spec-v6
		(8 bit mask)	
12	IPv4 Fragment	4 bit mask	RFC5575
13	IPv6 Flow	20 bit flow	ietf-idr-flow-spec-v6
14	Delimiter type	2 bytes	hao-idr-flowspec-nv03
MF-1	(Encapsulation type		
	VXLAN or NVGRE)		
15	VNID	24 bit VN	hao-idr-flowspec-nv03
MF-2	(virtual network ID)		
16	Flow ID	8 bit flow ID	hoa-idr-flowspec-nv03
MF-3	(NVGRE Flow ID)		
17	Segment ID		
18-25	Other packet ids		
	above MPLS		
29	MPLS LSP	TBD	not specified
MF-4	(label 20 bits,	Label stack	
	EXP (3 bits), S Bit		
	TTL (8 bits)		

30	Ethernet type	2 bytes	ietf-idr-flowspec-l2vpn
31	Source MAC	MAC address	ietf-idr-flowspec-l2vpn
32	Destination MAC	MAC Address	ietf-idr-flowspec-l2vpn
33	DSAP in LLC	1 octet	ietf-idr-flowspec-l2vpn
34	SSAP in LLC	1 octet	ietf-idr-flowspec-l2vpn
35	Control in LLC	1 octet	ietf-idr-flowspec-l2vpn
36	SNAP	5 octet	ietf-idr-flowspec-l2vpn
37	VLAN ID	1 or 2 bytes	ietf-idr-flowspec-l2vpn
38	VLAN COS	3 bit COS	ietf-idr-flowspec-l2vpn
39	Inner VLAN ID	1 or 2 bytes	ietf-idr-flowspec-l2vpn
40	Inner VLAN COS	1 or 2 bytes	ietf-idr-flowspec-l2vpn
41	Interface (Group ID, intf id)	TBD	not specified
42	Time		
65	FCFS matches		non-standard actions

Figure 12

5.3.2. FCFS Flow Specification Match Condition Filter Interaction

[RFC5575] allowed for non-IETF standardized Flow Specification filters and extended community actions. The beginning order of precedence for non-IETF standardized FCFS BGP Flow specification match filters is 65. The network management yang modules SHOULD store the BGP Flow Specification match type byte for both IETF Standardized BGP Flow Specification Match Filters, FCFS BGP BGP Flow Specification Match filters.

5.4. Flow Specification Actions and Action Precedence

Some BGP Flow Specification actions can conflict with other BGP Flow specification Actions. It will be the duty of each action specification to indicate how it interacts with the deafult precedence in Table 9 in figure 13 and the potential conflicts (shown in table 6 figure 9).

Table 9 provides the default precedence for actions for the minimal subset. All Standards actions have precedence overall FCFS actions incoded in BGP Extended Communities.

```
      Table 9 - Action Precedence and Conflicts between Actions
    +-----+----------------------------------------------------+
    |order| Action                                             |
    +=====+====================================================+
    |  1  | Alternate NLRI Validation (ROA, and future ROA) (FA7)|
    |  2  | Traffic Rate in bytes (0x8006)                     |
    |  3  | Traffic Rate in packets (FA1)                      |
    |  4  | Traffic Action (0x8007)  (T or S bit)              |
    |  5  | Redirect to IP-VPN  (0x8008, 0x8108, 0x8208)       |
    |     | 0x8008: 2 byte AS RD|                              |
    |     | 0x8108: 4 byte IP RD|                              |
    |     | 0x8208: 4 byte AS RD|                              |
    |  6  |  Redirect to IP Tunnel (FA3)                       |
    |  7  | VLAN action (FM4)                                  |
    |  8  | TPID action (FM5)                                  |
    |  9  | Label Action (FM6)                                 |
    | 10  | Interface set (FM8a)                               |
    | 11  | packet-ECA policy interaction                      |
    |     |  0 = BGP Flow-Specification (BGP FS) only          |
    |     |  1 = ACL + BGP FS                                  |
    |     |  2 = I2RS FB-RIB + BGP FS                          |
    |     |  3 = ACL + I2RS FB-FIB + BGP FS                    |
    |     |  4 = Config FB-RIB + BGP FS                        |
    |     |  5 = ACL + config FB-RIB + BGP FS                  |
    |     |  6 = Config FB-RIB + I2RS FB-RIB + BGP FS          |
    |     |  7 = ACL + config FB-FIB + I2RS                    |
    |12-64| Reserved for other standards actions               |
    |     |                                                    |
    |65+  | FCFS actions                                       |
    +=====+====================================================+
       Figure 13
```

5.4.1. FCFS Extended Communities with BGP Flow Specification Actions

[RFC7153]allows for FCFS (First Come First Serve) allocation of BGP transitive types. If an action is specified in the FCFS registry, the default precedence is after all standardized BGP Flow Specification actions(action 65+). The BGP Flow Specification Yang models should store the Extended Community value for the FCFS based Flow Specification action. If the precedence ordering has been changed by the FCFS, this should be stored in the configuration of BGP Flow Specification and in the operational state.

5.5. Precedence with other packet ECA policies

The BGP Flow Specification policy is currently handled as part of the route selection process within BGP. Between BGP and other n-tuple packet ECA policies, the precedence policies is handled by the

operator-applied policies (which often have operator default) which assign order and preference of filters within within an order. The default assumption for BGP-FS is to assume the worst possible valid order if none is specified (e.g. 254 out of 255), and to assume the priority within that order as shown in table 10. BGP Flow Specification (BGP-FS) Flow Specification for 128.2/16 destination port 20 may conflict with the following:

a) I2RS Flow Specification for destination address 128.2/16 with destination port 12, and

b) ACL filter for 128.2/16 destination address 128.2/16 with destination port 12.

In summmary, the precedence is least dynamic in configuration to most dynamic received. However, a BGP-FS action may signal a remote operator applied priority for a set of routes that allows the filters to combine certain filters (see table 11).

Table 10 - Precedence within a single order

priority	Filter source
10	BGP Flow Specification received from peer
9	BGP Flow Specification from Peer + BGP-FS action
8	BGP Flow Specification configured on local peer that is installed and distributed
7	I2RS Flow Specification
5	policy routing packet ECA filters configured
4	ACLS configured
3	policy configured in general routing table (netmod-routing-cfg)

Figure 14

Table 11 - actions that combine packet ECA policy

order	Action
11	packet-ECA policy interaction action based 0 = BGP Flow-Specification (BGP FS) only 1 = ACL + BGP FS 2 = I2RS FB-RIB + BGP FS 3 = ACL + I2RS FB-FIB + BGP FS 4 = Config FB-RIB + BGP FS 5 = ACL + config FB-RIB + BGP FS 6 = Config FB-RIB + I2RS FB-RIB + BGP FS 7 = ACL + config FB-FIB + I2RS
12-64	Reserved for other standards actions
65+	FCFS actions

Figure 15

5.6. pros and cons of Minimal subset BGP-FS Additions (Option 1)

Pro - for Minimal subset (Option 1)

Version 1's basic mechanism for BGP Flow Specification has been tested. Additions can be added incrementally.

Con - for Minimal Subset (Option 1)

The current version 1 of the Flow Specification does not have ordering of packet ECA policy rules, flow specification filters, or flow specification actions other than the default precedence. Current implementations of BGP flow specification are finding this lack of ordering to cause operational difficulties.

6. BGP-FS-v2 (New NLRI and Wide Communities Approach)(option 2)

This section on minimal subset solution has:

summary of NLRI and extended community formats (xection 6.1)

security addition of ROA (section 6.2),

match filter list and precedence of match filters (section 6.3),

action list and precedence of actions(section 6.4),

conflict with other Packet-reception Event-MatchCondition-Action (ECA) policy (I2RS Filter-Based RIB and Policy-Based Routing (n-tuple forwarding)) (section 6.5),

pros-cons of this approach (section 6.6)

6.1. Format of New NLRI and Wide Communities

The format of the NLRI TLVs would be replaced with:

```
+-------------------------+
|length (2 octets)        |
+-------------------------+
| sub-TLVs (variable)     |
| +====================+  |
| | order (2 octets)   |  |
| +--------------------+  |
| | type (2 octets)    |  |
| +--------------------+  |
| | length (2 octets)  |  |
| +--------------------+  |
| | value (variable)   |  |
| |[multiples of       |  |
| | 2 octets]          |  |
| +====================+  |
+-------------------------+
```

Figure 16 - NRLI revision

The Actions for BGP Flow Specification will be defined as an BGP Flow Specification Action atom within BGP Wide communities where the atom is defined as shown in figure 17.

```
+---------------------------+
| order (2 octets)          |
+---------------------------+
| Action type (2 octets)    |
+---------------------------+
| Action length (2 octets) |
+---------------------------+
| Action Values (variable) |
| (multiples of 2 octets)  |
+---------------------------+
```

Wide Community Atom
figure 17

The BGP Flow Specification (BGP-FS) atom can be part of the Wide Community container (type 1) or the BGP Flow Specification Atom can be part of the BGP Flow Specification container (type 2) which will have:

```
+----------------------------+
| Source AS Number  (4 octets)|
+----------------------------+
| list of atoms (variable)    |
+----------------------------+
```

figure 18

6.2. security addition of ROA

The security for the ROA is required to be the first action (action order 1) for all actions. All additional BGP Security precede all other security additions in the ordering.

6.3. Match Filters and precedence

The precedence of the match filters is determined by the order. If two orders are the same, the precedence is dependent on the order specified in the table below.

6.3.1. Precedence in case of ties in order

Table 9: Flow Specification Match Filter Precedence Order

type#	Type Name	Match	Reference
1	Destination Prefix	IPv4 Prefix	RFC5575
		IPv6 Prefix	ietf-idr-flow-spec-v6
2	Source Prefix	IPv4 Prefix	RFC5575
		IPv6 Prefix	ietf-idr-flow-spec-v6
3	IP protocol	IPv4 Protocol	RFC5575
		number	
3	Next Header	IPv6 protocol	ietf-idr-flow-spec-v6
4	Port (source or	Port number	RFC5575
	destination port)		RFC5575
5	Source port	Port number	RFC5575
6	Destination port	Port number	RFC5575
7	ICMP type	ICMP type	RFC5575
8	ICMP code	ICMP code	RFC5575
9	TCP Flags	1 or 2 byte	RFC5575
		bitmask for	RFC5575
		TCP flags	
10	Packet length	# of bytes	RFC5575
	(for IP packet)		

11	DSCP	IPv4 DSCP	RFC5575
		(6 bit mask)	RFC5575
11	Traffic class	IPv6 traffic	ietf-idr-flow-spec-v6
		(8 bit mask)	
12	IPv4 Fragment	4 bit mask	RFC5575
13	IPv6 Flow	20 bit flow	ietf-idr-flow-spec-v6
14	Delimiter type	2 bytes	hao-idr-flowspec-nv03
MF-1	(Encapsulation type		
	VXLAN or NVGRE)		
15	VNID	24 bit VN	hao-idr-flowspec-nv03
MF-2	(virtual network ID)		
16	Flow ID	8 bit flow ID	hoa-idr-flowspec-nv03
MF-3	(NVGRE Flow ID)		
17	Segment ID		
18-25	Other packet ids		
	above MPLS		
29	MPLS LSP	TBD	not specified
MF-4	(label 20 bits,	Label stack	
	EXP (3 bits), S Bit		
	TTL (8 bits)		
30	Ethernet type	2 bytes	ietf-idr-flowspec-l2vpn
31	Source MAC	MAC address	ietf-idr-flowspec-l2vpn
32	Destination MAC	MAC Address	ietf-idr-flowspec-l2vpn
33	DSAP in LLC	1 octet	ietf-idr-flowspec-l2vpn
34	SSAP in LLC	1 octet	ietf-idr-flowspec-l2vpn
35	Control filed in		
	LLC|	1 octet	ietf-idr-flowspec-l2vpn
36	SNAP	5 octet	ietf-idr-flowspec-l2vpn
37	VLAN ID	1 or 2 bytes	ietf-idr-flowspec-l2vpn
38	VLAN COS	3 bit COS	ietf-idr-flowspec-l2vpn
39	Inner VLAN ID	1 or 2 bytes	ietf-idr-flowspec-l2vpn
40	Inner VLAN COS	1 or 2 bytes	ietf-idr-flowspec-l2vpn
41	Interface	TBD	not specified
	(Group ID, intf id)		
42	Time		
65	FCFS matches		non-standard actions

Figure 19

6.3.2. Precedence of filters among Routing Functions

As discussed in the minimum sub-set (Option 1 for BGP-FS), there needs to be a precedence between n-tuple packet ECA policies. This precedence is determined by policy rule order and a preference among policy rules with the same order. Match Condition order is defined by the BGP-FS Filter order, and within the match the action order is defined by the BGP-FS.

Precedence among policy rules from difference sources with the same order is commonly specified by operator-applied policies (which may be supplied by vendor defaults) where lower priority implies a better route. For example, a BGP Flow Specification Policy rule can be set to a priority of 150 where an static ACL policy might be set to a priority of 40. If the same two n-tuple packet ECA policies exist, then the lower priority rule within the the same order is selected to be active.

The operator-applied policy can change these priorities globally or for a specific route.

If any packet ECA related policy changes, then the BGP Flow specification must be re-evaluated per policy rule per order and priority.

6.3.3. Precedence for re-ordering Match Policy

Actions that change interact between levels of policy need to be defined in terms of policy actions in BGP Flow Specification. For example [I-D.litkowski-idr-flowspec-interfaceset] provides a definition of the following combination of filter rules between ACLs and BGP flow Specifications:

1. Forward if both ACL forward and BGP Flow Specification Forward

2. Drop if either ACL drops or BGP Flow Specification drops.

6.4. Actions and precedence of actions

The actions allowed for BGP are listed in Table 12 provides the default precedence for actions for the minimal subset. All Standards actions have precedence overall FCFS actions incoded in BGP Extended Communities. The default order for these actions are listed below. All drafts defining actions must deal with the conflicts between actions and the ordering (see section 4).

Table 10 - Action Precedence and Conflicts between Actions

order	Action
1	Alternate NLRI Validation (ROA, and future ROA) (FA7)
2	Alternate bgpsec validation
5	Traffic Rate in bytes (0x8006)
6	Traffic Rate in packets (FA1)
7	Traffic Action (0x8007) (T or S bit)
8	Extension to Traffic Actions
-10	
11	Redirect to IP-VPN (0x8008, 0x8108, 0x8208)
	0x8008: 2 byte AS RD|
	0x8108: 4 byte IP RD|
	0x8208: 4 byte AS RD|
12	Redirect to IP Tunnel (FA3)
13-20}	redirect actions (other)
21	VLAN action (FM4)
22	TPID action (FM5)
23	Label Action (FM6)
30	Interface set (FM8a)
40	packet-ECA policy interaction
	0 = BGP Flow-Specification (BGP FS) only
	1 = ACL + BGP FS
	2 = I2RS FB-RIB + BGP FS
	3 = ACL + I2RS FB-FIB + BGP FS
	4 = Config FB-RIB + BGP FS
	5 = ACL + config FB-RIB + BGP FS
	6 = Config FB-RIB + I2RS FB-RIB + BGP FS
	7 = ACL + config FB-FIB + I2RS
50	Time
51-64	Reserved for other standards actions
65+	FCFS actions

Figure 20

6.5. Pro-Con of BGP-FS-v2 (option 2}

Pro - for version 2

The current version 1 of the Flow Specification does not have ordering of packet ECA policy rules, flow specification filters, or flow specification actions other than the default precedence. Current implementations of BGP flow specification are finding this lack of ordering to cause operational difficulties.

Con - for version 2

Version 2 must be coded. It can either be a BGP attribute with the policy rules (NLRI filters and actions) inside such as described in [I-D.li-idr-flowspec-rpd] or it can be a combination of a new BGP Flow Specification version 2 NLRI + Wide Community actions (with ordering).

(Additional comments will be added here)

7. Flow Specification Yang models

The Flow Specification Yang models have configuration and operational state. BGP Flow Specification (BGP-FS) configuration have local configuration for BGP-FS and locally configured BGP-FS policy rules. Operational state has three components:

1. Local node's BGP-FS Operational Configuration installed (if supported)

2. BGP Flow specification rules received from peers,

3. BGP Flow Specfication match statistics

Comparison of the the BGP local configuration for BGP-FS policy rules with the BGP-FS policy rules, is aided by common yang definitions between these two functions. Comparison of the BGP-FS Policy rules (locally configured or received) with I2RS Filter-Based RIB (FB-RIB), packet-ECA policy, ACL policy rules, and routing table policy rules requires is aided by common yang definitions between packet-ECA filtesr.

This section compares BGP Flow Specification yang model in [I-D.wu-idr-flowspec-yang-cfg] and the I2RS FB-RIB data model is described in [I-D.hares-i2rs-fb-rib-data-model] which uses the packet reception ECA policy data model found in [I-D.hares-i2rs-pkt-eca-data-model]. A comparison of the policy structures is given in table 8, and the operation status model is given in table 9. These models are similar. It would be helpful to use a common yang definitions found in [I-D.hares-i2rs-pkt-eca-data-model].

The packet reception ECA policy data model is also used to describe configured packet reception filter RIBs which (aka Policy Routing) described in (draft-hares-rtgwg-fb-rib-00.txt).

Table 11 - comparison Yang Model Local Configuraoin

```
+-------------+---------------------+----------------------+
| component    | BGP Flow Spec       | I2RS FB-RIB  +        |
|              | Yang                | Packet-ECA Yang       |
+==============+=====================+======================+
|Policy        |flowspec-policy*     |group* [group-name]    |
| +-name       | [policy-name]       |                       |
| +-vrf        |+-rw vrf-name        | +-rw vrf-name         |
| +-AFI        |+-rw address-family  | +-rw address-famil    |
| +-rules      |+-rw flowspec-rule*  | +-rw group-rule-list  |
|              ||  [rule-name]       | | [rule-name]         |
|  +-rule-name ||+-rw rule-name      | |+-rw rule-name       |
|  +-rule-order||+-rw traffic-filters| |+-rw rule-order      |
|              ||+-rw traffic-actions| +-rw eca-rules        |
|              |                     | | [order-id rule-name]|
|              |                     | | +-rw installer      |
|              |                     | | +-rw eca-matches    |
|              |                     | | +-rw eca-qos-actions|
|              |                     | | +-rw eca-fwd-actions|
+--------------+---------------------+----------------------+
```

figure 21 - Comparison of Yang modules (Config state)

Note:The Yang "traffic-filters" found are the same as eca-matches found in [I-D.wu-idr-flowspec-yang-cfg] are the same filters found in [I-D.hares-i2rs-pkt-eca-data-model]. The "traffic actions" found in [I-D.wu-idr-flowspec-yang-cfg] can be broken into modify actions and forwarding actions as [I-D.hares-i2rs-pkt-eca-data-model] does.

Table 12 - comparison of Yang operational state

```
+------------+----------------------+------------------------+
| component  | BGP Flow Spec        | I2RS FB-RIB            |
|            | Yang                 | Packet-ECA Yang        |
+============+======================+========================+
|opstate     |flowspec-state        |ietf-fb-ribs-oper-status |
| +-rib      |+-ro flowspec-rib     |+-ro fb-rib-oper-status* |
|            |  |                   |  +-ro fb-rib-name       |
|  +-groups  |  |                   |  +-ro group-status      |
|  +-rules   |  +-ro flowspec-entry*|  +-ro rules_opstate     |
|    [index] |    [index]           |  [rule-order, rule-name]|
|            |                      |  |                      |
|statistics  |                      |  |                      |
| +-rules    |+-ro flowspec-stats*  |  +-ro rules_opstats     |
|            |                      |  [rule-order, rule-name]|
|            | +-ro vrf-name        |                         |
|            | +-ro address-family  |                         |
|            | +-ro flowspec-rule-  |                         |
|            | |    stats           |                         |
|            | |                    |                         |
|            | |+-ro traffic-filters|                         |
|            | |+-ro traffic-action |                         |
|            | |+-ro classified-pkts|  | +--ro pkts-match     |
|            | |                    |  | +--ro pkts-modified  |
|            | |+-ro drop-pkts      |  | +--ro pkts-dropped   |
|            | |+-ro drop-bytes     |  | +--ro bytes-dropped  |
|            |                      |  | +--ro pkts-forwarded |
|            |                      |  | +--ro bytes-forwarded|
+------------+----------------------+------------------------+
```

figure 22 - Comparison of Yang Models (Operation State)

8. IANA Considerations

This section complies with [RFC7153]

TBD. There are a lot of assignments which will be filled in after the initial review of the technology.

9. Security Considerations

The new BGP Validation described in section 4.1 with the ROA improves on [RFC5575] security by improving the validation of the originating AS having permissions to send Flow specfcation for a prefix. The validation of the path attributes and/or path requires the BGPSEC [I-D.ietf-sidr-bgpsec-protocol]. [I-D.eddy-idr-flowspec-exp] contains suggestions on how to implement this with flow specification, but at this time the authors consider the technology

described in [I-D.eddy-idr-flowspec-exp] so this draft does not suggest mandating it. However, it encourages the develop of such work that pairs BGP Flow Specification with BGPSEC protocol. When this work matures, this specification or BGP Flow Specification version 2 should implement it.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://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, <http://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, <http://www.rfc-editor.org/info/rfc4360>.

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

[RFC4761] Kompella, K., Ed. and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, DOI 10.17487/RFC4761, January 2007, <http://www.rfc-editor.org/info/rfc4761>.

[RFC4762] Lasserre, M., Ed. and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, DOI 10.17487/RFC4762, January 2007, <http://www.rfc-editor.org/info/rfc4762>.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <http://www.rfc-editor.org/info/rfc5226>.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009, <http://www.rfc-editor.org/info/rfc5575>.

[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, <http://www.rfc-editor.org/info/rfc6241>.

[RFC6482] Lepinski, M., Kent, S., and D. Kong, "A Profile for Route Origin Authorizations (ROAs)", RFC 6482, DOI 10.17487/RFC6482, February 2012, <http://www.rfc-editor.org/info/rfc6482>.

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

[RFC7223] Bjorklund, M., "A YANG Data Model for Interface Management", RFC 7223, DOI 10.17487/RFC7223, May 2014, <http://www.rfc-editor.org/info/rfc7223>.

[RFC7674] Haas, J., Ed., "Clarification of the Flowspec Redirect Extended Community", RFC 7674, DOI 10.17487/RFC7674, October 2015, <http://www.rfc-editor.org/info/rfc7674>.

10.2. Informative References

[I-D.eddy-idr-flowspec-exp]
Eddy, W., Dailey, J., and G. Clark, "Experimental BGP Flow Specification Extensions", draft-eddy-idr-flowspec-exp-00 (work in progress), August 2015.

[I-D.eddy-idr-flowspec-packet-rate]
Eddy, W., Dailey, J., and G. Clark, "BGP Flow Specification Packet-Rate Action", draft-eddy-idr-flowspec-packet-rate-00 (work in progress), November 2015.

[I-D.hao-idr-flowspec-nvo3]
Weiguo, H., Zhuang, S., Li, Z., and R. Gu, "Dissemination of Flow Specification Rules for NVO3", draft-hao-idr-flowspec-nvo3-03 (work in progress), December 2015.

[I-D.hao-idr-flowspec-redirect-tunnel]
Weiguo, H., Li, Z., and L. Yong, "BGP Flow-Spec Redirect to Tunnel action", draft-hao-idr-flowspec-redirect-tunnel-00 (work in progress), October 2015.

[I-D.hares-i2rs-fb-rib-data-model]
Hares, S., Kini, S., Dunbar, L., Krishnan, R., Bogdanovic, D., and R. White, "Filter-Based RIB Data Model", draft-hares-i2rs-fb-rib-data-model-02 (work in progress), February 2016.

[I-D.hares-i2rs-pkt-eca-data-model]
Hares, S., Wu, Q., and R. White, "Filter-Based Packet Forwarding ECA Policy", draft-hares-i2rs-pkt-eca-data-model-02 (work in progress), February 2016.

[I-D.ietf-i2rs-architecture]
Atlas, A., Halpern, J., Hares, S., Ward, D., and T. Nadeau, "An Architecture for the Interface to the Routing System", draft-ietf-i2rs-architecture-13 (work in progress), February 2016.

[I-D.ietf-i2rs-ephemeral-state]
Haas, J. and S. Hares, "I2RS Ephemeral State Requirements", draft-ietf-i2rs-ephemeral-state-02 (work in progress), September 2015.

[I-D.ietf-idr-flow-spec-v6]
Raszuk, R., Pithawala, B., McPherson, D., and A. Andy, "Dissemination of Flow Specification Rules for IPv6", draft-ietf-idr-flow-spec-v6-06 (work in progress), November 2014.

[I-D.ietf-idr-flowspec-l2vpn]
Weiguo, H., Litkowski, S., and S. Zhuang, "Dissemination of Flow Specification Rules for L2 VPN", draft-ietf-idr-flowspec-l2vpn-03 (work in progress), November 2015.

[I-D.ietf-idr-wide-bgp-communities]
Raszuk, R., Haas, J., Lange, A., Amante, S., Decraene, B., Jakma, P., and R. Steenbergen, "Wide BGP Communities Attribute", draft-ietf-idr-wide-bgp-communities-01 (work in progress), November 2015.

[I-D.ietf-netconf-restconf]
Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", draft-ietf-netconf-restconf-09 (work in progress), December 2015.

[I-D.ietf-netmod-acl-model]
Bogdanovic, D., Koushik, K., Huang, L., and D. Blair, "Network Access Control List (ACL) YANG Data Model", draft-ietf-netmod-acl-model-06 (work in progress), December 2015.

[I-D.ietf-netmod-opstate-reqs]
Watsen, K. and T. Nadeau, "Terminology and Requirements for Enhanced Handling of Operational State", draft-ietf-netmod-opstate-reqs-04 (work in progress), January 2016.

[I-D.ietf-netmod-routing-cfg]
Lhotka, L. and A. Lindem, "A YANG Data Model for Routing Management", draft-ietf-netmod-routing-cfg-20 (work in progress), October 2015.

[I-D.ietf-sidr-bgpsec-protocol]
Lepinski, M., "BGPsec Protocol Specification", draft-ietf-sidr-bgpsec-protocol-14 (work in progress), December 2015.

[I-D.kini-i2rs-fb-rib-info-model]
Kini, S., Hares, S., Dunbar, L., Ghanwani, A., Krishnan, R., Bogdanovic, D., and R. White, "Filter-Based RIB Information Model", draft-kini-i2rs-fb-rib-info-model-03 (work in progress), February 2016.

[I-D.li-idr-flowspec-rpd]
Li, Z., Ou, L., Luo, Y., Lu, S., Zhuang, S., and N. Wu, "BGP FlowSpec Extensions for Routing Policy Distribution (RPD)", draft-li-idr-flowspec-rpd-01 (work in progress), October 2015.

[I-D.liang-idr-bgp-flowspec-label]
You, J., Raszuk, R., and d. danma@cisco.com, "Carrying Label Information for BGP FlowSpec", draft-liang-idr-bgp-flowspec-label-01 (work in progress), September 2015.

[I-D.liang-idr-bgp-flowspec-time]
You, J. and S. Zhuang, "BGP FlowSpec with Time Constraints", draft-liang-idr-bgp-flowspec-time-00 (work in progress), October 2015.

[I-D.litkowski-idr-flowspec-interfaceset]
Litkowski, S., Simpson, A., Patel, K., and J. Haas, "Applying BGP flowspec rules on a specific interface set", draft-litkowski-idr-flowspec-interfaceset-03 (work in progress), December 2015.

[I-D.rosen-idr-tunnel-encaps]
Rosen, E., Patel, K., and G. Velde, "Using the BGP Tunnel Encapsulation Attribute without the BGP Encapsulation SAFI", draft-rosen-idr-tunnel-encaps-03 (work in progress), August 2015.

[I-D.vandevelde-idr-flowspec-path-redirect]
Velde, G., Henderickx, W., and K. Patel, "Flowspec Indirection-id Redirect", draft-vandevelde-idr-flowspec-path-redirect-01 (work in progress), January 2016.

[I-D.wu-idr-flowspec-yang-cfg]
Wu, N., Zhuang, S., and A. Choudhary, "A YANG Data Model for Flow Specification", draft-wu-idr-flowspec-yang-cfg-02 (work in progress), October 2015.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, DOI 10.17487/RFC4303, December 2005, <http://www.rfc-editor.org/info/rfc4303>.

[RFC6074] Rosen, E., Davie, B., Radoaca, V., and W. Luo, "Provisioning, Auto-Discovery, and Signaling in Layer 2 Virtual Private Networks (L2VPNs)", RFC 6074, DOI 10.17487/RFC6074, January 2011, <http://www.rfc-editor.org/info/rfc6074>.

[RFC6483] Huston, G. and G. Michaelson, "Validation of Route Origination Using the Resource Certificate Public Key Infrastructure (PKI) and Route Origin Authorizations (ROAs)", RFC 6483, DOI 10.17487/RFC6483, February 2012, <http://www.rfc-editor.org/info/rfc6483>.

Author's Address

Susan Hares
Huawei
7453 Hickory Hill
Saline, MI 48176
USA

Email: shares@ndzh.com