IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 27, 2016

S. Hares
Huawei

June 25, 2016

BGP Flow Specification Version 2
draft-hares-idr-flowspec-v2-00.txt

Abstract

BGP flow specification version 1 (RFC5575) describes the distribution of traffic filter policy (traffic filters and actions) which are distributed via BGP to BGP peers. Three applications utilize this traffic filter policy: (1) mitigation of Denial of Service (DoS), (2) enabling of traffic filtering in BGP/MPLS VPNS, and (3)centralized traffic control for networks with SDN or NFV controllers. Application of centralized traffic utilizing BGP Flow Specification traffic filters may need user-ordered filters rather than RFC5575's strict ordering of filters and defined ordering of actions.

This document proposes a new BGP Flow specification version 2 that supports user-order of filters and actions plus allowing more actions

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction 2
1.1. RFC5575 vs. NETCONF/RESTCONF/I2RS Flow Filters 4
2. Definitions 6
2.1. Definitions and Acronyms 6
2.2. RFC 2119 language 6
3. Dissemination of BGP Flow Specification version 2 NLRI and Wide Communities 6
3.1. Encoding of BGP-FS v2 Filters 7
3.2. Encoding of BGP-FS v2 Actions 7
3.3. Required NLRI Validation 8
4. Optional Security Additions 8
4.1. BGP FS v2 and BGPSEC 9
4.2. BGP FS v2 with with ROA 9
4.3. Revise Flow Specification Security for centralized Server 9
5. IANA Considerations 10
6. Security Considerations 11
7. References 11
7.1. Normative References 11
7.2. Informative References 13
Author’s Address 14

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 were the match-condition is defined in the BGP NLRI, and the action is defined either by the default condition (accept traffic) or actions defined in Extended BGP Communiites values [RFC4360].

The initial set of policy [RFC5575] and [RFC7674] for this policy includes 12 types of match filters encoded in two application specific AFI/SAFIs for the IPv4 AFI.

IP traffic: AFI:1, SAFI, 133;

BGP/MPLS VPN AFI:1 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.

This document describes distribution of two new BGP Flow Specification NLRI (2 AFI/SAFI pairs) that allow user-ordered list of traffic match filters, and user-ordered traffic match actions encoded in BGP Wide Communities.

- section 2 - Definitions,
- section 3 - Rules for dissemination of Flow Specification v2,
- section 4 - Optional Security,
- section 5 - IANA considerations,
- section 6 - security considerations.

The rest of this section provides background on BGP Flow Specification filters interaction with I2RS Filter-Based RIBs carried by NETCONF/RESTCONF protocol. Figure 1 below is a logial description of BGP Flow Specification rules that combine filters in BGP NLRI with actions in BGP Extended communities.

```
        +-----------------------------+
        | 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      |  |        | |      | |(subtype| |        ||      |
    +--------+  +--------+ +------+ +--------+ +--------++------+
```

*1 match operator may be complex.

Figure 1: BGP Flow Specification Policy

BGP Flow Specification (BGP-FS) ([RFC5575] and [I-D.raszuk-idr-rfc5575bis]) describes how to distribute the BGP Flow Specification policy as BGP routes which are locally configured on the originating BGP peer. Like BGP routes, if the BGP peer session drops then BGP Flow Specification routes are dropped. [RFC5575] and [I-D.raszuk-idr-rfc5575bis] do not indicate how the BGP Flow Specification policy is installed in the kernel.

1.1. RFC5575 vs. NETCONF/RESTCONF/I2RS Flow Filters

[RFC5575] describes the dissemination of flow specification rules policy is similar to the the statically configured Filter-Based RIB described in [I-D.ietf-i2rs-fb-rib-data-model], and the I2RS Filter-Based RIB ([I-D.ietf-i2rs-fb-rib-info-model], [I-D.ietf-i2rs-fb-rib-data-model], [I-D.ietf-i2rs-pkt-eca-data-model]). These FB-RIBs start on the reception of a packet using match filters to match frames (L2) or packet data (L3/L4/Application), and perform actions as shown in figure 2.

```
       +-----------+     +------------+
       |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

[I-D.ietf-i2rs-fb-rib-data-model] suggests that the storage of BGP Flow Specification routes in the kernel should utilize the same format as the statically configured FB-RIB and the I2RS ephemeral FB-RIB so that these traffic filters may be compared. This draft also proposes that precedence between these three sources of filters in the kernel (statically configured, I2RS ephemeral, and BGP ephemeral routes) can either set by local policy or defaults. If it is set by defaults [I-D.ietf-i2rs-fb-rib-data-model] suggests the default precedence between static, I2RS, and BGP-FS installed filters is:

o static FB-RIB -highest precedence (wins all ties)

o I2RS FB-RIB - middle preference (wins over BGP-FS originated routes, loses to static FB-RIB),

- o BGP-FS installed Filters - lows preference (loses to static and I2RS FB-RIB)

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].

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

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.

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. Dissemination of BGP Flow Specification version 2 NLRI and Wide Communities

The BGP Flow Specification version 2 (BGP-FS v2) uses an NRLI with the format for AFI/SAFI (SAFI = TBD) for IP flow, and AFI/SAFI for BGP/MPLS (SAFI = TBD). This NLRI information is encoded using MP_READ_NRI and MP_UNREACH_NLRI attributes defined in [RFC4760]. Whenever the corresponding application does not require Next-HOP information, this shall be encoded as zero-octet length Next Hop in the MP_REAC_NLRI and ignored upon receipt.

Implementatinos wishing to exchange flow specificastion rules MUST use BGP’s Capability Advertisement facility to exchange the Multiprotocol Extension Capability Code (Code 1) as defined in [RFC4760].

3.1. Encoding of BGP-FS v2 Filters

The AFI/SAFI NLRI for BGP Flow Specification has the format

```
     +------------------------+
     |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

where:

- o length - is the length of the NLRI,
- o Sub-TLVs contain a user-ordered set of filter components as defined in [RFC5575] and [I-D.raszuk-idr-rfc5575bis]. The ranges are defined as: standard BGP Flow Specification filters (types 0x01 - 0x3FFFF), and vendor specific filters (types 0x4ffff to 0x7FFFF) with type values 0x8000 to 0xFFFFFFFF reserved for future use. Each sub-tlv has an length of 2 octets, and a variable length value (in multiples of 2 octets).

Filters are process in the order specified by the user. If multiple filters exist for the same order, the strict filter ordering defined in [RFC5575] and [I-D.raszuk-idr-rfc5575bis] will be used for the filters with the same value for user order.

3.2. Encoding of BGP-FS v2 Actions

The BGP-FS version 2 actions are passed in a Wide Community [I-D.ietf-idr-wide-bgp-communities] atom with the following format.

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

Wide Community Atom
figure 17

where:

- o Action type (2 octets) - is the type of action. These actions can be standardized (0x0001 - 0x3ffff), vendor specific (0x40000-0x7FFFF), or reserved (0x0, 0x80000-0xFFFFFFFF).

- o Action length - length of actions including variable field,

- o Action values - value of actions (variable) defined in individual definitions.

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

3.3. Required NLRI Validation

Same as [RFC5575] and [I-D.raszuk-idr-rfc5575bis].

4. Optional Security Additions

This section discusses the optional BGP Security additions for BGP-FS v2: BGPSEC [I-D.ietf-sidr-bgpsec-protocol], ROA [RFC6482] and revised security for flow specification distributed from a centralized server within an AS [I-D.ietf-idr-bgp-flowspec-oid]. These optional security parameters can be applied per BGP peer.

4.1. BGP FS v2 and BGPSEC

[RFC5575] does not require BGP Flow specifications to be passed BGPSEC [I-D.ietf-sidr-bgpsec-protocol]. BGP FS v2 can be passed in BGPSEC, but it is not required.

4.2. BGP FS v2 with with ROA

BGP-FS v2 can utilize ROAs in the validation. If BGP-FS v2 is used with BGPSEC and ROA, the first thing is to vaildate 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:

- 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.

4.3. Revise Flow Specification Security for centralized Server

The distribution of Flow Specifications from a centralized server supports mitigation of DoS attacks. [I-D.ietf-idr-bgp-flowspec-oid] suggests the following redefined procedure for validation for this case:

A route is valid if the following conditions holds true:

- o The originator of the flow specification matches the originator of the best-match unicast route for the destination prefix embedded in the flow specification.

- o The AS_PATH and AS4_PATH attribute of the flow specification are empty (on originating AS)

- o The AS_PATH and AS4_PATH attribute of the flow specification does not contain AS_SET and AS_SEQUENCE segments (on originating AS with AS Confederation)

This reduced validation mechanism can be used for BGP-FS v2 within a single domain.

5. IANA Considerations

This section complies with [RFC7153]

This document requests:

SAFI be defined for IPv4 (AFI = 1), IPv6 (AFI=2), L2VPN (AFI=25) for BGP-FS

SAFI be defined for BGP/MPLS IPv4 (AFI = 1), IPv6 (AFI=2), L2VPN (AFI=25) for BGP-FS

Registry be created for BGP-FS V2 filter component types with the following ranges:

0x00 - reserved

0x01 - 0x3FFFF - standards action

0x40000- 0x7FFFF - vendor specific filters

0x80000 -0xFFFFFFFF - reserved

0x80000 -0xFFFFFFFF - reserved

Registry be created for BGP-FS v2 action types with the following ranges:

0x0 - reserved

0x01 - 0x3ffff - standards action

0x40000 - 0x7ffff - vendor actions

0x80000 - 0xFFFFFFF - reserved

6. Security Considerations

The use of ROA improves on [RFC5575] to check the route orgination is valid can improve the validation sequence for a multiple-AS environment. The use of BGPSEC [I-D.ietf-sidr-bgpsec-protocol] 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 [I-D.ietf-idr-bgp-flowspec-oid] can provide adequate validation for distribution of flow specification within an 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.

7. References

7.1. Normative References

[I-D.ietf-idr-bgp-flowspec-oid]
Uttaro, J., Filsfils, C., Smith, D., Alcaide, J., and P. Mohapatra, "Revised Validation Procedure for BGP Flow Specifications", draft-ietf-idr-bgp-flowspec-oid-03 (work in progress), March 2016.

[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-02 (work in progress), May 2016.

[I-D.ietf-sidr-bgpsec-protocol]
Lepinski, M. and K. Sriram, "BGPsec Protocol Specification", draft-ietf-sidr-bgpsec-protocol-17 (work in progress), June 2016.

[I-D.raszuk-idr-rfc5575bis]
Raszuk, R., McPherson, D., Mauch, J., Greene, B., and S. Hares, "Dissemination of Flow Specification Rules", draft-raszuk-idr-rfc5575bis-00 (work in progress), June 2016.

[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>.

7.2. Informative References

[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-15 (work in progress), April 2016.

[I-D.ietf-i2rs-ephemeral-state] Haas, J. and S. Hares, "I2RS Ephemeral State Requirements", draft-ietf-i2rs-ephemeral-state-10 (work in progress), June 2016.

[I-D.ietf-i2rs-fb-rib-data-model] Hares, S., Kini, S., Dunbar, L., Krishnan, R., Bogdanovic, D., and R. White, "Filter-Based RIB Data Model", draft-ietf-i2rs-fb-rib-data-model-00 (work in progress), June 2016.

[I-D.ietf-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-ietf-i2rs-fb-rib-info-model-00 (work in progress), June 2016.

[I-D.ietf-i2rs-pkt-eca-data-model] Hares, S., Wu, Q., and R. White, "Filter-Based Packet Forwarding ECA Policy", draft-ietf-i2rs-pkt-eca-data-model-00 (work in progress), June 2016.

[I-D.ietf-netconf-restconf] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", draft-ietf-netconf-restconf-13 (work in progress), April 2016.

[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-07 (work in progress), March 2016.

[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