

INTERNET-DRAFT
Intended Status: Proposed Standard

W. Hao
Huawei Technologies
D. Eastlake
Futurewei Technologies
J. Uttaro
AT&T
S. Litkowski
Orange Business Service
S. Zhuang
Huawei Technologies"
July 8, 2019

Expires: January 7, 2020

BGP Dissemination of L2VPN Flow Specification Rules
draft-ietf-idr-flowspec-l2vpn-11

Abstract

This document defines a BGP flow-spec extension to disseminate L2 VPN Ethernet traffic filtering rules. SAFI=134 in [draft-ietf-idr-rfc5575bis](#) is redefined for this purpose. A new subset of component types and extended community also are defined. A new subset of component types and new extended community also are defined.

Status of This Document

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Distribution of this document is unlimited. Comments should be sent to the authors or the TRILL Working Group mailing list <dnsext@ietf.org>.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/1id-abstracts.html>. The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

Table of Contents

1. Introduction.....	3
1.1 Terminology.....	4
2. Layer 2 Flow Specification encoding in BGP.....	5
3. Ethernet Flow Specification encoding in BGP.....	6
3.1 Order of Traffic Filtering Rules.....	8
4. Ethernet Flow Specification Traffic Actions.....	9
4.1 VLAN-action.....	9
4.2 TPID-action.....	11
5. IANA Considerations.....	12
6. Security Considerations.....	13
7. Acknowledgements.....	13
8. Contributors.....	13
Normative References.....	14
Informative References.....	14
Authors' Addresses.....	15

1. Introduction

BGP Flow-spec is an extension to BGP that supports the dissemination of traffic flow specification rules. It leverages the BGP Control Plane to simplify the distribution of ACLs. Using this extension new filter rules can be injected to all BGP peers simultaneously without changing router configuration. The typical application of BGP Flow-spec is to automate the distribution of traffic filter lists to routers for DDOS mitigation, access control, etc.

[RFC5575bis] defines a new BGP Network Layer Reachability Information (NLRI) format used to distribute traffic flow specification rules. NLRI (AFI=1, SAFI=133) is for IPv4 unicast filtering. NLRI (AFI=1, SAFI=134) is for BGP/MPLS VPN filtering. The Flow specification match part only includes L3/L4 information like source/destination prefix, protocol, ports, and etc., so traffic flows can only be selectively filtered based on L3/L4 information.

Layer 2 Virtual Private Networks (L2VPNs) have already been deployed in an increasing number of networks today. In an L2VPN network, we also have requirements to deploy BGP Flow-spec to mitigate DDoS attack traffic. Within an L2VPN network, both IP and non-IP Ethernet traffic maybe exist. For IP traffic filtering, the Flow specification rules defined in [RFC5575bis] which include match criteria and actions can still be used, flow specification rules received via new NLRI format apply only to traffic that belongs to the VPN instance(s) in which it is imported. For non-IP Ethernet traffic filtering, Layer 2 related information like source/destination MAC and VLAN should be considered. But the flow specification match criteria defined in [RFC5575bis] only include layer 3 and layer 4 IP information, not layer 2 Ethernet information.

There are different kinds of L2VPN networks like EVPN [RFC7432], BGP VPLS [RFC4761], LDP VPLS [RFC4762] and border gateway protocol (BGP) auto discovery [RFC6074]. Because the flow-spec feature relies on BGP protocol to distribute traffic filtering rules, it can only be incrementally deployed in those L2VPN networks where BGP has already been used for auto discovery and/or signaling purposes such as BGP-based VPLS [RFC4761], EVPN and LDP-based VPLS [RFC4762] with BGP auto-discovery [RFC6074].

This draft proposes a new subset of flow-spec component types and an extended community to support L2VPN flow-spec application. The flow-spec rules can be enforced on all border routers or on some interface sets of the border routers. SAFI=134 in [RFC5575bis] is redefined for dissemination of traffic filtering information in an L2VPN environment.

1.1 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [[RFC2119](#)] [[RFC8174](#)] when, and only when, they appear in all capitals, as shown here.

2. Layer 2 Flow Specification encoding in BGP

[RFC5575bis] defines SAFI 133 and SAFI 134 for "dissemination of IPv4 flow specification rules" and "dissemination of VPNv4 flow specification rules" respectively. [[I-D.ietf-idr-flow-spec-v6](#)] redefines the [[RFC5575bis](#)] SAFIs in order to make them applicable to both IPv4 and IPv6 applications. This document will further redefine the SAFI 134 in order to make them applicable to L2VPN applications.

The following changes are defined:

"SAFI 134 for dissemination of L3VPN flow specification rules" to now be defined as "SAFI 134 for dissemination of VPN flow specification rules"

For SAFI 134 the indication to which address family it is referring to will be recognized by AFI value (AFI=1 for VPNv4, AFI=2 VPNv6 and AFI=25 for L2VPN). Such modification is fully backwards compatible with existing implementation and production deployments.

For SAFI 134 the indication to which address family it is referring to will be recognized by AFI value (AFI=1 for VPNv4, AFI=2 VPNv6 and AFI=25 for L2VPN). Such modification is fully backwards compatible with existing implementation and production deployments.

3. Ethernet Flow Specification encoding in BGP

The NLRI format for this address family consists of a fixed-length Route Distinguisher field (8 bytes) followed by a flow specification, following the encoding defined in this document. The NLRI length field includes both the 8 bytes of the Route Distinguisher as well as the subsequent flow specification.

Flow specification rules received via this NLRI apply only to traffic that belongs to the VPN instance(s) into which it is imported. Flow rules are accepted by default when received from remote PE routers.

Besides the component types defined in [[RFC5575bis](#)] and [[I-D.ietf-idr-flow-spec-v6](#)], this document specifies the following additional component types for L2 VPN Ethernet traffic filtering:

Type 14 - Ethernet Type

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match two-octet field. Values are encoded as 2-byte quantities. Ethernet II framing defines the two-octet Ethernet Type (EtherType) field in an Ethernet frame, preceded by destination and source MAC addresses, that identifies an upper layer protocol encapsulating the frame data.

Type 15 - Source MAC

Encoding: <type (1 octet), MAC Address length (1 octet), MAC Address>

Defines the source MAC Address to match.

Type 16 - Destination MAC

Encoding: <type (1 octet), MAC Address length (1 octet), MAC Address>

Defines the destination MAC Address to match.

Type 17 - DSAP(Destination Service Access Point) in LLC

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match the 1-octet DSAP in the 802.2 LLC (Logical Link Control Header). Values are encoded as 1-byte quantities. The operation field is encoded as a

Type 18 - SSAP(Source Service Access Point) in LLC

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match the 1-octet SSAP in the 802.2 LLC. Values are encoded as 1-byte

quantities.

Type 19 - Control field in LLC

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match 1-octet control field in the 802.2 LLC. Values are encoded as 1-byte quantities.

Type 20 - SNAP

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match 5-octet SNAP (Sub-Network Access Protocol) field. Values are encoded as 5-byte quantities.

Type 21 - VLAN ID

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match VLAN ID. Values are encoded as 2-byte quantities, where the four most significant bits are zero and the 12 least significant bits contain the VLAN value.

In the virtual local-area network (VLAN) stacking case, the VLAN ID is the outer VLAN ID.

Type 22 - VLAN COS

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match 3-bit VLAN COS fields [802.1Q]. Values are encoded using a single byte, where the five most significant bits are zero and the three least significant bits contain the VLAN COS value.

In the virtual local-area network (VLAN) stacking case, the VLAN COS is outer VLAN COS.

Type 23 - Inner VLAN ID

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match the inner VLAN ID using for virtual local-area network (VLAN) stacking or Q in Q use. Values are encoded as 2-byte quantities, where the four most significant bits are zero and the 12 least significant bits contain the VLAN value.

In single VLAN case, this component type MUST NOT be used.

Type 24 - Inner VLAN COS

Encoding: <type (1 octet), length (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match 3-bit inner VLAN COS fields [802.1Q] using for virtual local-area network (VLAN) stacking or Q in Q use. Values are encoded using a single byte, where the five most significant bits are zero and the three least significant bits contain the VLAN COS value.

In single VLAN case, the component type MUST NOT be used.

3.1 Order of Traffic Filtering Rules

The original definition for the order of traffic filtering rules can be reused with new consideration for the MAC Address offset. As long as the offsets are equal, the comparison is the same, retaining longest-prefix-match semantics. If the offsets are not equal, the lowest offset has precedence, as this flow matches the most significant bit.

Pseudocode:

```
flow_rule_L2_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;
        }

        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
        }
    }
    return EQUAL;
}
```

}

4. Ethernet Flow Specification Traffic Actions

The default action for a layer 2 traffic filtering flow specification is to accept traffic that matches that particular rule. The following extended community values per [\[RFC5575bis\]](#) can be used to specify particular actions in an L2 VPN network:

type	extended community	encoding
0x8006	traffic-rate	2-byte as#, 4-byte float
0x8007	traffic-action	bitmask
0x8008	redirect	6-byte Route Target
0x8009	traffic-marking	DSCP value

Redirect: The action should be redefined to allow the traffic to be redirected to a MAC or IP VRF routing instance that lists the specified route-target in its import policy.

Besides the above extended communities, this document also specifies the following BGP extended communities for Ethernet flows to extend [\[RFC5575bis\]](#):

type	extended community	encoding
TBD1	VLAN-action	bitmask
TBD2	TPID-action	bitmask

4.1 VLAN-action

The VLAN-action extended community, as shown in the diagram below, consists of 6 bytes that include 2 action Flags, two VLAN IDs, and the associating COS value. The action Flags fields are further divided into two parts which correspond to the first action and the second action respectively. Bit 0 to bit 7 give the first action while bit 8 to bit 15 give the second action. The bits of PO, PU, SW, RI and RO in each part represent the action of Pop, Push, Swap, Rewrite inner VLAN and Rewrite outer VLAN respectively. Through this method, more complicated actions also can be represented in a single VLAN-action extended community, such as SwapPop, PushSwap, etc. For example, SwapPop action is the sequence of two actions, the first action is Swap and the second action is Pop.


```

    0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|P01|PU1|SW1|RI1|R01| Resv      |P02|PU2|SW2|RI2|R02| Resv      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| VLAN ID1                                |COS1          |R1  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| VLAN ID2                                |COS2          |R2  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

P01: Pop action. If the P01 flag is one, it indicates the outmost VLAN should be removed.

PU1: Push action. If PU1 is one, it indicates VLAN ID1 will be added, the associated COS is COS1.

SW1: Swap action. If the SW1 flag is one, it indicates the outer VLAN and inner VLAN should be swapped.

P02: Pop action. If the P02 flag is one, it indicates the outmost VLAN should be removed.

PU2: Push action. If PU2 is one, it indicates VLAN ID2 will be added, the associated COS is COS2.

SW2: Swap action. If the SW2 flag is one, it indicates the outer VLAN and inner VLAN should be swapped.

RI1 and RI2: Rewrite inner VLAN action. If the RI flag is one, it indicates the inner VLAN should be replaced by a new VLAN where the new VLAN is VLAN ID1 and the associated COS is COS1. If the VLAN ID1 is 0, the action is to only modify the COS value of inner VLAN.

R01 and R02: Rewrite outer VLAN action. If the R0 flag is one, it indicates the outer VLAN should be replaced by a new VLAN where the new VLAN is VLAN ID and the associated COS is COS2. If the VLAN ID2 is 0, the action is to only modify the COS value of outer VLAN.

Resv, R1, and R2: Reserved for future use. MUST be sent as zero and ignored on receipt.

Giving an example below: if the action of PUSH Inner VLAN 10 with COS value 5 and Outer VLAN 20 with COS value 6 is needed, the format of the VLAN-action extended community is as follows:


```

    0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| 10                                     | 1 | 0 | 1 | 0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| 20                                     | 1 | 1 | 0 | 0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

4.2 TPID-action

The TPID-action extended community consists of 6 bytes which includes the fields of action Flags, TPID1 and TPID2.

```

    0                                     15
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| TI | T0 |                               Resv                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     TP ID1                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     TP ID2                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

TI: Mapping inner TP ID action. If the TI flag is one, it indicates the inner TP ID should be replaced by a new TP ID, the new TP ID is TP ID1.

T0: Mapping outer TP ID action. If the T0 flag is one, it indicates the outer TP ID should be replaced by a new TP ID, the new TP ID is TP ID2.

Resv: Reserved for future use. MUST be sent as zero and ignored on receipt.

5. IANA Considerations

IANA is requested to change the description for SAFI 134 [[RFC5575bis](#)] to read as follows more general description and to change the reference for it to [this document]:

134 VPN dissemination of flow specification rules

IANA is requested to allocate 11 new values in the Flow-Spec Component Type registry as follows:

type	RFC or Draft	discription
14	This draft	Ethernet Type
15	This draft	Source MAC
16	This draft	Destination MAC
17	This draft	DSAP in LLC
18	This draft	SSAP in LLC
19	This draft	Control field in LLC
20	This draft	SNAP
21	This draft	VLAN ID
22	This draft	VLAN COS
23	This draft	Inner VLAN ID
24	This draft	Inner VLAN COS

IANA is requested to update the reference for the following assignment in the "BGP Extended Communities Type - extended, transitive" registry:

Type value	Name	Reference
0x080A	Flow spec VLAN action	[this document]
0x080B	Flow spec TPID action	[this document]

6. Security Considerations

No new security issues are introduced to the BGP protocol by this specification.

7. Acknowledgements

The authors wish to acknowledge the important contributions of Hannes Gredler, Xiaohu Xu, Zhenbin Li, Lucy Yong, and Feng Dong.

8. Contributors

Qiandeng Liang
Huawei Technologies
101 Software Avenue, Yuhuatai District
Nanjing 210012
China

Email: liangqiandeng@huawei.com

Normative References

- [I-D.ietf-idr-flow-spec-v6] McPherson, D., Raszuk, R., Pithawala, B., akarch@cisco.com, a., and S. Hares, "Dissemination of Flow Specification Rules for IPv6", [draft-ietf-idr-flow-spec-v6-09](#) (work in progress), November 2017.
- [RFC5575bis] Hares, S., Loibl, C., Raszuk, R., McPherson, D., Bacher, M., "Dissemination of Flow Specification Rules", [draft-ietf-idr-rfc5575bis-17](#), Work in progress, June 2019.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [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, <<https://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, <<https://www.rfc-editor.org/info/rfc4762>>.
- [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, <<https://www.rfc-editor.org/info/rfc6074>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

Informative References

- [RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", [RFC 7432](#), DOI 10.17487/RFC7432, February 2015, <<https://www.rfc-editor.org/info/rfc7432>>.

Authors' Addresses

Weiguo Hao
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

Email: haoweiguo@huawei.com

Donald E. Eastlake, 3rd
Futurewei Technologies
1424 Pro Shop Court
Davenport, FL 33896
USA

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

James Uttaro
AT&T

Email: uttaro@att.com

Stephane Litkowski
Orange Business Service

Email: stephane.litkowski@orange.com

Shunwan Zhuang
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: zhuangshunwan@huawei.com

Copyright, Disclaimer, and Additional IPR Provisions

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

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

