IDR

Internet Draft
Intended status: Standards Track

Expires: November 2016

W. Hao
Q. Liang
Huawei
Jim Uttaro
AT&T
S. Litkowski
Orange Business Service
S. Zhuang
Huawei

May 17, 2016

Dissemination of Flow Specification Rules for L2 VPN
draft-ietf-idr-flowspec-l2vpn-04.txt

Abstract

This document defines BGP flow-spec extension for Ethernet traffic filtering in L2 VPN network. SAFI=134 in [RFC5575] is redefined for dissemination traffic filtering information in an L2VPN environment. A new subset of component types and extended community also are defined.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), 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

BGP Flow-spec is an extension to BGP that allows for the dissemination of traffic flow specification rules. It leverages the BGP Control Plane to simplify the distribution of ACLs, 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.

RFC5575 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 L2VPN network, we also have requirement to deploy BGP Flow-spec to mitigate DDoS attack

traffic. Within L2VPN network, both IP and non-IP Ethernet traffic maybe exist. For IP traffic filtering, the Flow specification rules defined in [RFC5575] 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 RFC5575 only include layer 3 and layer 4 IP information, layer 2 Ethernet information haven't been included.

There are different kinds of L2VPN networks like EVPN [EVPN], BGP VPLS [RFC4761], LDP VPLS [RFC4762] and border gateway protocol (BGP) auto discovery [RFC 6074]. Because the flow-spec feature relies on BGP protocol to distribute traffic filtering rules, so 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 [4761], EVPN and LDP-based VPLS [4762] with BGP auto-discovery [6074].

This draft proposes a new subset of component types and 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 [RFC5575] is redefined for dissemination traffic filtering information in an L2VPN environment.

2. Layer 2 Flow Specification encoding in BGP

The [RFC5575] defines SAFI 133 and SAFI 134 for ''dissemination of IPv4 flow specification rules'' and ''dissemination of VPNv4 flow specification rules'' respectively. [draft-ietf-idr-flow-spec-v6-06] redefines the [RFC5575] 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.

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 shall include 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) in which it is imported. Flow rules are accepted by default, when received from remote PE routers.

Besides the component types defined in [RFC5575] and [draft-ietf-idr-flow-spec-v6-06], this document proposes the following additional component types for L2VPN Ethernet traffic filtering:

Type 14 - Ethernet Type

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

Defines a list of {operation, value} pairs used to match 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 'Numeric operator' defined in [RFC5575].

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

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

Defines a list of {operation, value} pairs used to match 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), [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), [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), [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 virtual local-area network (VLAN) stacking case, the VLAN ID is outer VLAN ID.

Type 22 - VLAN COS

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

Defines a list of {operation, value} pairs used to match 3-bit VLAN COS fields [802.1p]. 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 virtual local-area network (VLAN) stacking case, the VLAN COS is outer VLAN COS.

Type 23 - Inner VLAN ID

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

Defines a list of {operation, value} pairs used to match inner VLAN ID using for virtual local-area network (VLAN) stacking or Q in Q case. 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, the component type MUST not be used.

Type 24 - Inner VLAN COS

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

Defines a list of {operation, value} pairs used to match 3-bit inner VLAN COS fields [802.1p] using for virtual local-area network (VLAN) stacking or Q in Q case. 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 RFC5575 can be used to specify particular actions in L2VPN network:

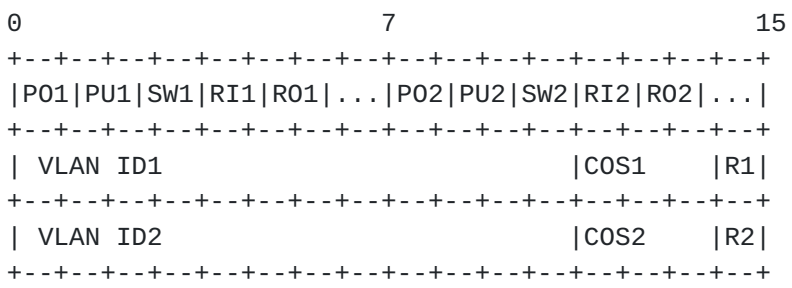
type	extended community	encoding
0x8006	traffic-rate	2-byte as#, 4-byte float
0x8007	traffic-action	bitmask
0x8008	redirect	6-byte Route Target
0x8009	traffic-marking	DSCP value

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 proposes the following BGP extended communities specifications for Ethernet flow to extend [RFC5575]:

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

VLAN-action: The VLAN-action extended community consists of 6 bytes which include the fields of 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 belong to the first action part while bit 8 to bit 15 belong to the second part. 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 extend community, such as SwapPop, PushSwap, etc. For example, SwapPop action is the concatenation of two actions, the first action is Swap and the second action is Pop.



PO1: Pop action. If the PO1 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.

PO2: Pop action. If the PO2 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.

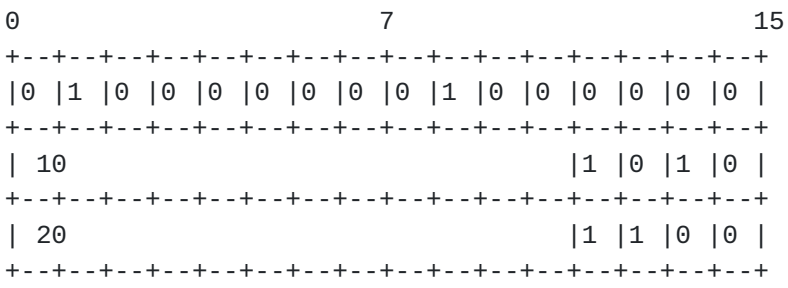
SW2: Swap action. If the SW2 flag is one it indicates the outer VLAN and inner VLAN should be swapped.

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, the new VLAN is VLAN ID1, the associated COS is COS1. If the VLAN ID1 is 0, the action is to only modify the COS value of inner VLAN.

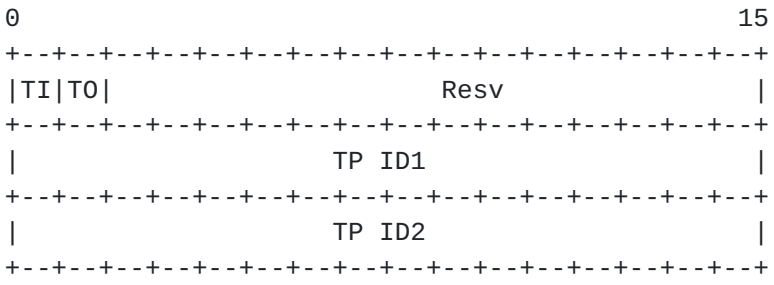
RO1 and RO2: Rewrite outer VLAN action. If the RO flag is one it indicates the outer VLAN should be replaced by a new VLAN, the new VLAN is VLAN ID2, the associated COS is COS2. If the VLAN ID2 is 0, the action is to only modify the COS value of outer VLAN.

R1 and R2: Reserved for future use.

Giving an example, 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:



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



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.

TO: Mapping outer TP ID action. If the TO 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.

5. Security Considerations

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

6. IANA Considerations

IANA is requested to rename currently defined SAFI 134 per [RFC5575] to read:

134 VPN dissemination of flow specification rules

IANA is requested to create and maintain a new registry for "Flow spec L2VPN Component Types". For completeness, the types defined in [RFC5575] and [draft-ietf-idr-flow-spec-v6-06] also are listed here.

type	RFC or Draft	discription
1	RFC5575	Destination Prefix
1	draft-ietf-idr-flow-spec-v6-06	Destination IPv6 Prefix
2	RFC5575	Source Prefix
2	draft-ietf-idr-flow-spec-v6-06	Source IPv6 Prefix
3	RFC5575	IP Protocol
3	draft-ietf-idr-flow-spec-v6-06	Next Header
4	RFC5575	Port
5	RFC5575	Destination port
6	RFC5575	Source port
7	RFC5575	ICMP type
8	RFC5575	ICMP code
9	RFC5575	TCP flags
10	RFC5575	Packet length
11	RFC5575	DSCP
12	RFC5575	Fragment
13	draft-ietf-idr-flow-spec-v6-06	Flow Label
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]

6.1. Normative References

[1] [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[2] [RFC5575] P. Marques, N. Sheth, R. Raszuk, B. Greene, J.Mauch, D. McPherson, "Dissemination of Flow Specification Rules", RFC 5575, August 2009.

[3] [RFC4761] K. Kompella, Ed., Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC4761, January 2007.

[4] [RFC4762] M. Lasserre, Ed., V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC4762, January 2007.

[5] [RFC6074] E. Rosen, B. Davie, V. Radoaca, "Provisioning, Auto-Discovery, and Signaling in Layer 2 Virtual Private Networks (L2VPNs)", RFC6074, January 2011.

6.2. Informative References

[1] [EVPN] Sajassi et al., "BGP MPLS Based Ethernet VPN", draft-ietf-l2vpn-evpn-07.txt, work in progress, May, 2014.

[2] [IEEE 802.1p] Javin, et.al. "IEEE 802.1p: LAN Layer 2 QoS/CoS Protocol for Traffic Prioritization", 2012-02-15

7. Acknowledgments

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

Authors' Addresses

Weiguo Hao
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China
Email: haoweiguo@huawei.com

Qiandeng Liang
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China
Email: liangqiandeng@huawei.com

James Uttaro
AT&T
Email: uttaro@att.com

Stephane Litkowski
Orange
Email: stephane.litkowski@orange.com

Shunwan Zhuang
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China
Email: zhuangshunwan@huawei.com