Dissemination of Flow Specification Rules for L2 VPN
draft-hao-idr-flowspec-evpn-02.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.

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not be created, and it may not be published except as an Internet-Draft.

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not be created, except to publish it as an RFC and to translate it into languages other than English.

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."
1. Introduction ................................................ 2
2. Layer 2 Flow Specification encoding in BGP.................. 3
3. Ethernet Flow Specification encoding in BGP............... 4
4. Ethernet Flow Specification Traffic Actions............... 7
5. Security Considerations.................................... 7
6. IANA Considerations ........................................ 7
   6.1. Normative References.................................... 8
   6.2. Informative References.................................. 9
7. Acknowledgments ............................................ 9
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.

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 is 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. 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 – Source MAC

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

Defines a list of (operation(op), value) pairs used to match source MAC. Values are encoded as 6-byte quantities. The operation field is encoded as ’’Numeric operator’’ defined in [RFC5575].

Type 15 – Destination MAC

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

Defines a list of (operation, value) pairs used to match destination MAC. Values are encoded as 6-byte quantities.
Type 16 - 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 EtherType field in an Ethernet frame, preceded by destination and source MAC addresses, that identifies an upper layer protocol encapsulating the frame data.

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.

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 1- or 2-byte quantities.

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 1- or 2-byte quantities.

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.
4. Ethernet Flow Specification Traffic Actions

<table>
<thead>
<tr>
<th>type</th>
<th>extended community</th>
<th>encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x8006</td>
<td>traffic-rate</td>
<td>2-byte as#, 4-byte float</td>
</tr>
<tr>
<td>0x8007</td>
<td>traffic-action</td>
<td>bitmask</td>
</tr>
<tr>
<td>0x8008</td>
<td>redirect</td>
<td>6-byte Route Target</td>
</tr>
<tr>
<td>0x8009</td>
<td>traffic-marking</td>
<td>DSCP value</td>
</tr>
</tbody>
</table>

Besides to support the above extended communities per RFC5575, this document also proposes the following BGP extended communities specifications for Ethernet flow to extend [RFC5575):

<table>
<thead>
<tr>
<th>type</th>
<th>extended community</th>
<th>encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x800A</td>
<td>VLAN COS marking</td>
<td>COS value</td>
</tr>
</tbody>
</table>

The VLAN COS marking extended community instructs a system to modify the COS bits of a transiting Ethernet packet to the corresponding value. This extended community is encoded as a sequence of 5 zero bytes followed by the VLAN COS value encoded in the 3 least significant bits of 6th byte.

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

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.
<table>
<thead>
<tr>
<th>type</th>
<th>RFC or Draft</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RFC5575</td>
<td>Destination Prefix</td>
</tr>
<tr>
<td>2</td>
<td>RFC5575</td>
<td>Source Prefix</td>
</tr>
<tr>
<td>3</td>
<td>RFC5575</td>
<td>IP Protocol</td>
</tr>
<tr>
<td>4</td>
<td>RFC5575</td>
<td>Next Header</td>
</tr>
<tr>
<td>5</td>
<td>RFC5575</td>
<td>Port</td>
</tr>
<tr>
<td>6</td>
<td>RFC5575</td>
<td>Destination port</td>
</tr>
<tr>
<td>7</td>
<td>RFC5575</td>
<td>Source port</td>
</tr>
<tr>
<td>8</td>
<td>RFC5575</td>
<td>ICMP type</td>
</tr>
<tr>
<td>9</td>
<td>RFC5575</td>
<td>ICMP code</td>
</tr>
<tr>
<td>10</td>
<td>RFC5575</td>
<td>TCP flags</td>
</tr>
<tr>
<td>11</td>
<td>RFC5575</td>
<td>Packet length</td>
</tr>
<tr>
<td>12</td>
<td>RFC5575</td>
<td>DSCP</td>
</tr>
<tr>
<td>13</td>
<td>draft-ietf-idr-flow-spec-v6-06</td>
<td>Flow Label</td>
</tr>
<tr>
<td>14</td>
<td>This draft</td>
<td>Source MAC</td>
</tr>
<tr>
<td>15</td>
<td>This draft</td>
<td>Destination MAC</td>
</tr>
<tr>
<td>16</td>
<td>This draft</td>
<td>Ethernet Type</td>
</tr>
<tr>
<td>17</td>
<td>This draft</td>
<td>DSAP in LLC</td>
</tr>
<tr>
<td>18</td>
<td>This draft</td>
<td>SSAP in LLC</td>
</tr>
<tr>
<td>19</td>
<td>This draft</td>
<td>Control field in LLC</td>
</tr>
<tr>
<td>20</td>
<td>This draft</td>
<td>SNAP</td>
</tr>
<tr>
<td>21</td>
<td>This draft</td>
<td>VLAN ID</td>
</tr>
<tr>
<td>22</td>
<td>This draft</td>
<td>VLAN COS</td>
</tr>
<tr>
<td>23</td>
<td>This draft</td>
<td>Inner VLAN ID</td>
</tr>
<tr>
<td>24</td>
<td>This draft</td>
<td>Inner VLAN COS</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Type value</th>
<th>Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x080A</td>
<td>Flow spec VLAN COS marking</td>
<td>[this document]</td>
</tr>
</tbody>
</table>

6.1. Normative References

6.2. Informative References


7. Acknowledgments

The authors wish to acknowledge the important contributions of Xiaohu Xu, 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

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

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

Stephane Litkowski
Orange
stephane.litkowski@orange.com