IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 19, 2018

D. McPherson
Verisign, Inc.
R. Raszuk, Ed.
Bloomberg LP
B. Pithawala
Individual
A. Karch
Cisco Systems
S. Hares, Ed.
Huawei

November 15, 2017

Dissemination of Flow Specification Rules for IPv6
draft-ietf-idr-flow-spec-v6-09.txt

Abstract

Dissemination of Flow Specification Rules [RFC5575] provides a protocol extension for propagation of traffic flow information for the purpose of rate limiting or filtering. The [RFC5575] specifies those extensions for IPv4 protocol data packets.

This specification extends the current [RFC5575] and defines changes to the original document in order to make it also usable and applicable to IPv6 data packets.

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). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on May 19, 2018.

Table of Contents

1. Introduction

The growing amount of IPv6 traffic in private and public networks requires the extension of tools used in the IPv4 only networks to be also capable of supporting IPv6 data packets.

In this document authors analyze the differences of IPv6 [RFC2460] flows description from those of traditional IPv4 packets and propose subset of new encoding formats to enable Dissemination of Flow Specification Rules [RFC5575] for IPv6.

This specification should be treated as an extension of base [RFC5575] specification and not its replacement. It only defines the delta changes required to support IPv6 while all other definitions and operation mechanisms of Dissemination of Flow Specification Rules will remain in the main specification and will not be repeated here.

2. IPv6 Flow Specification encoding in BGP

The [RFC5575] defines a new SAFIs (133 for IPv4) and (134 for VPNv4) applications in order to carry corresponding to each such application flow specification.

This document will redefine the [RFC5575] SAFIs in order to make them AFI specific and applicable to both IPv4 and IPv6 applications.

The following changes are defined:

"SAFI 133 for IPv4 dissemination of flow specification rules" to now be defined as "SAFI 133 for dissemination of unicast flow specification rules"

"SAFI 134 for VPNv4 dissemination of flow specification rules" to now be defined as "SAFI 134 for dissemination of L3VPN flow specification rules"

For both SAFIs the indication to which address family they are referring to will be recognized by AFI value (AFI=1 for IPv4 or VPNv4, AFI=2 for IPv6 and VPNv6 respectively). Such modification is fully backwards compatible with existing implementation and production deployments.

It needs to be observed that such choice of proposed encoding is compatible with filter validation against routing reachability information as described in section 6 of RFC5575. Validation tables will now be performed according to the following rules.

Flow specification received over AFI/SAFI=1/133 will be validated against routing reachability received over AFI/SAFI=1/1

Flow specification received over AFI/SAFI=1/134 will be validated against routing reachability received over AFI/SAFI=1/128

Flow specification received over AFI/SAFI=2/133 will be validated against routing reachability received over AFI/SAFI=2/1

Flow specification received over AFI/SAFI=2/134 will be validated against routing reachability received over AFI/SAFI=2/128

3. IPv6 Flow Specification types changes

The following component types are redefined or added for the purpose of accommodating new IPv6 header encoding. Unless otherwise stated all other types as defined in [RFC5575] apply to IPv6 packets as is.

Type 1 - Destination IPv6 Prefix

Encoding: <type (1 octet), prefix length (1 octet), prefix offset (1 octet), prefix>

Function: Defines the destination prefix to match. Prefix offset has been defined to allow for flexible matching on part of the IPv6 address where we want to skip (don't care) of N first bits of the address. This can be especially useful where part of the IPv6 address consists of an embedded IPv4 address and matching needs to happen only on the embedded IPv4 address. The encoded prefix contains enough octets for the bits used in matching (length minus offset bits).

Type 2 - Source IPv6 Prefix

Encoding: <type (1 octet), prefix length (1 octet), prefix offset (1 octet), prefix>

Function: Defines the source prefix to match. Prefix offset has been defined to allow for flexible matching on part of the IPv6 address where we want to skip (don't care) of N first bits of the address. This can be especially useful where part of the IPv6 address consists of an embedded IPv4 address and matching needs to happen only on the embedded IPv4 address. The encoded prefix contains enough octets for the bits used in matching (length minus offset bits)

Type 3 - Next Header

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

Function: Contains a set of {operator, value} pairs that are used to match the last Next Header value octet in IPv6 packets. The operator byte is encoded as specified in component type 3 of [RFC5575].

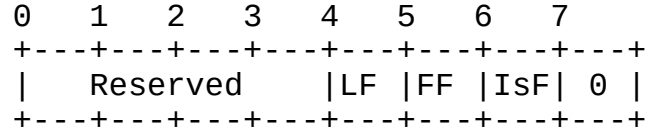
Note: While IPv6 allows for more then one Next Header field in the packet the main goal of Type 3 flow specification component is to match on the subsequent IP protocol value. Therefor the definition is limited to match only on last Next Header field in the packet.

Type 12 - Fragment

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

Uses bitmask operand format defined above. Bit-7 is not used and MUST be 0 to provide backwards-compatibility with the definition in [RFC5575]

Bitmast operand format:



Bitmask values:

+ Bit 6 - Is a fragment (IsF)

+ Bit 5 - First fragment (FF)

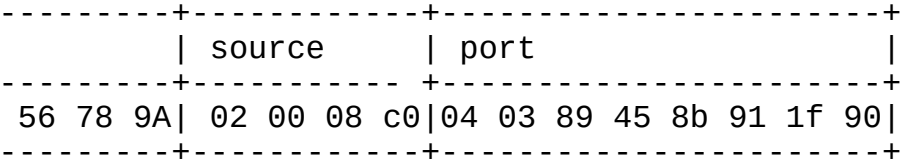
+ Bit 4 - Last fragment (LF)

Type 13 - Flow Label (New type)

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

Function: Contains a set of {operator, value} pairs that are used to match the 20-bit Flow Label field [RFC2460]. The operator byte is encoded as specified in the component type 3 of [RFC5575]. Values are encoded as 1-, 2-, or 4- byte quantities.

The following example demonstrates the new prefix encoding for: "all packets to ::1234:5678:9A00:0/64-104 from 192::/8 and port {range [137, 139] or 8080}". In the destination prefix, "80-" represents the prefix offset of 80 bits. In this exmaple, the 0 offset is omitted from the printed source prefix.



3.1. Order of Traffic Filtering Rules

The orignal definition for the order of traffic filtering rules can be reused with new consideration for the IPv6 prefix 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_v6_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) == IPV6_DESTINATION || IPV6_SOURCE) {
        // offset not equal, lowest offset has precedence
        // offset equal ...
        common_len = MIN(prefix_length(comp1), prefix_length(comp2));
        cmp = prefix_compare(comp1, comp2, offset, common_len);
        // 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. IPv6 Flow Specification Traffic Filtering Action changes

One of the traffic filtering actions which can be expressed by BGP extended community is defined in [RFC5575] as traffic-marking. Another traffic filtering action defined in [RFC5575] as a BGP extended community is redirect. To allow an IPv6 address specific route-target, a new traffic action IPv6 address specific extended community is provided.

Therefore, for the purpose of making it compatible with IPv6 header action expressed by presence of the extended community the following text in [RFC5575] has been modified to read:

Traffic Marking (0x8009): The traffic marking extended community instructs a system to modify first 6 bits of Traffic Class field as (recommended by [RFC2474]) of a transiting IPv6 packet to the corresponding value. This extended community is encoded as a sequence of 42 zero bits followed by the 6 bits overwriting DSCP portion of Traffic Class value.

Redirect-IPv6 (0x800B): redirect IPv6 address specific extended community allows the traffic to be redirected to a VRF routing instance that lists the specified IPv6 address specific route-target in its import policy. If several local instances match this criteria, the choice between them is a local matter (for example, the instance with the lowest Route Distinguisher value can be elected). This extended community uses the same encoding as the IPv6 address specific Route Target extended community [RFC5701].

5. Security Considerations

No new security issues are introduced to the BGP protocol by this specification over the security concerins in [RFC5575]

6. IANA Considerations

This section complies with [RFC7153]

IANA is requested to rename currently defined SAFI 133 and SAFI 134 per [RFC5575] to read:

133	Dissemination of flow specification rules
134	L3VPN dissemination of flow specification rules

IANA is requested to create and maintain a new registry entitled: "Flow Spec IPv6 Component Types". The initial values are:

Type	Description	RFC
Type 1	Destination IPv6 Prefix	[this draft]
Type 2	Source IPv6 Prefix	[this draft]
Type 3	Next Header	[this draft]
Type 4	Port	[this draft]
Type 5	Destination port	[this draft]
Type 6	Source port	[this draft]
Type 7	ICMP type	[this draft]
Type 8	ICMP code	[this draft]
Type 9	TCP flags	[this draft]
Type 10	Packet length	[this draft]
Type 11	DSCP	[this draft]
Type 12	Fragment	[this draft]
Type 13	Flow Label	[this draft]

7. Acknowledgements

Authors would like to thank Pedro Marques, Hannes Gredler and Bruno Rijsman, Brian Carpenter, and Thomas Mangin for their valuable input.

8. References

8.1. 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, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <https://www.rfc-editor.org/info/rfc2460>.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998, <https://www.rfc-editor.org/info/rfc2474>.

[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, <https://www.rfc-editor.org/info/rfc4271>.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <https://www.rfc-editor.org/info/rfc5492>.

[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, <https://www.rfc-editor.org/info/rfc5575>.

[RFC5701] Rekhter, Y., "IPv6 Address Specific BGP Extended Community Attribute", RFC 5701, DOI 10.17487/RFC5701, November 2009, <https://www.rfc-editor.org/info/rfc5701>.

[RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme, "IPv6 Flow Label Specification", RFC 6437, DOI 10.17487/RFC6437, November 2011, <https://www.rfc-editor.org/info/rfc6437>.

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

8.2. Informative References

[RFC5095] Abley, J., Savola, P., and G. Neville-Neil, "Deprecation of Type 0 Routing Headers in IPv6", RFC 5095, DOI 10.17487/RFC5095, December 2007, <https://www.rfc-editor.org/info/rfc5095>.

Authors' Addresses

Danny McPherson
Verisign, Inc.

Email: dmcpherson@verisign.com

Robert Raszuk (editor)
Bloomberg LP
731 Lexington Ave
New York City, NY 10022
USA

Email: robert@raszuk.net

Burjiz Pithawala
Individual

Email: burjizp@gmail.com

Andy Karch
Cisco Systems
170 West Tasman Drive
San Jose, CA 95134
USA

Email: akarch@cisco.com

Susan Hares (editor)
Huawei
7453 Hickory Hill
Saline, MI 48176
USA

Email: shares@ndzh.com