Network Working Group
Internet Draft
Intended status: Standards Track
Nov 26, 2012
Expires: May 26, 2013

James Uttaro
AT&T
Matthieu Texier
Arbor Networks
David Smith
Pradosh Mohapatra
Cisco Systems
Wim Henderickx
Adam Simpson
Alcatel-Lucent

BGP Flow-Spec Extended Community for Traffic Redirect to IP Next Hop

draft-simpson-idr-flowspec-redirect-02.txt

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/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on May 26, 2013.

Copyright Notice

Copyright (c) 2012 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.

Abstract

Flow-spec is an extension to BGP that allows for the dissemination of traffic flow specification rules. This has many possible applications but the primary one for many network operators is the distribution of traffic filtering actions for DDoS mitigation. The flow-spec standard [RFC 5575] defines a redirect-to-VRF action for policy-based forwarding but this mechanism can be difficult to use, particularly in networks without L3 VPNs.

This draft proposes a new redirect-to-IP flow-spec action that provides a simpler method of policy-based forwarding. This action is indicated by the presence of a new BGP extended community in the flow-spec route. Many routers already support a redirect-to-IP filter action and, in this case, the only new functionality implied by this draft is the ability to signal the action using flow-spec.

Table of Contents

1. Introduction

Flow-spec is an extension to BGP that allows for the dissemination of traffic flow specification rules. This has many possible applications but the primary one for many network operators is the distribution of traffic filtering actions for DDoS mitigation.

Every flow-spec route is effectively a rule, consisting of a matching part (encoded in the NLRI field) and an action part (encoded as a BGP extended community). The flow-spec standard [RFC 5575] defines widely-used filter actions such as discard and rate limit; it also defines a redirect-to-VRF action for policy-based forwarding. Using the redirect-to-VRF action for redirecting traffic towards an alternate destination is useful for DDoS mitigation but it can be complex and cumbersome, particularly in networks without L3 VPNs.

This draft proposes a new redirect-to-IP flow-spec action that provides a simpler method of policy-based forwarding. This action is indicated by the presence of a new BGP extended community in the flow-spec route. Many routers already support a redirect-to-IP filter action and, in this case, the only new functionality implied by this draft is the ability to signal the action using flow-spec.

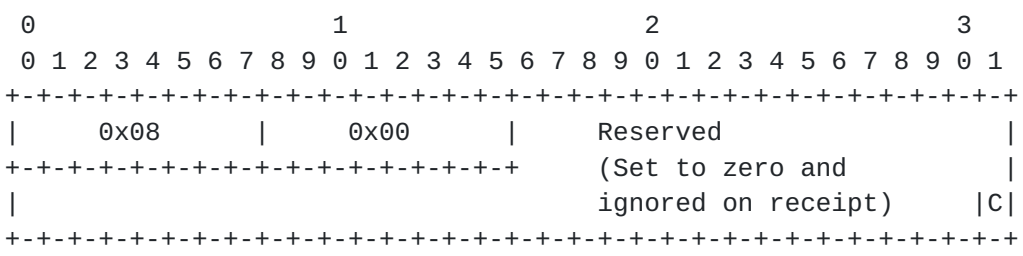
2. Terminology

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 [RFC-2119].

3. Redirect to IP Extended Community

This document proposes a new BGP extended community called "Flow spec redirect/mirror to IP next-hop" with type value 0x0800 (assigned from the "BGP Extended Communities Type - extended, transitive" registry). The new extended community, simply called "redirect to IP" in the remainder of this document, can be added to any UPDATE message announcing the reachability of one or more flow-spec NLRI. The encoding of the attribute is shown in Figure 1. In the 6 bytes of data after the 2-byte type value the least-significant bit is the 'C' (copy) bit. If 'C' is equal to 1 the originator of the flow-spec route is requesting a mirror action: routers that install this flow-spec route should create a copy of every matching packet and forward the copies towards a specified next-hop address while still forwarding the original packets normally (i.e. based on longest-prefix-match forwarding table lookups). If 'C' is equal to 0 the originator of the flow-spec route

is requesting a simple redirect action: routers that install this flow-spec route should forward the matching packets (the original versions, not copies) towards a new next-hop address. All bits other than the 'C' bit in the 6-byte data portion of the extended community should be set to 0 by the originating BGP speaker and ignored by receiving BGP speakers.



Flow-spec Redirect/Mirror to IP Next-hop Extended Community

Figure 1

The redirect-to-IP extended community is valid with any other set of flow-spec extended communities except if that set includes a redirect-to-VRF extended community (type 0x8008) and in that case the redirect-to-IP extended community should be ignored.

When a BGP speaker receives an UPDATE message with the redirect-to-IP extended community it is expected to create a traffic filtering rule for every flow-spec NLRI in the message that has this path as its best path. The filter entry matches the IP packets described in the NLRI field and redirects them (C=0) or copies them (C=1) towards the IPv4 or IPv6 address specified in the 'Network Address of Next-Hop' field of the associated MP_REACH_NLRI. More specifically: if an IPv4 [or IPv6] packet with destination address D that is normally forwarded to a next-hop A matches a filter entry of the type described above it MUST instead be redirected (C=0) or mirrored (C=1) to next-hop B, where B is found by FIB lookup of the IPv4 [or IPv6] address contained in the MP_REACH_NLRI next-hop field (i.e. a longest-prefix-match lookup). Signaling and applying constraints beyond longest-prefix-match on the types of interfaces or tunnels that can be used as the redirection next-hop B are not precluded by this specification but are nevertheless outside its scope.

If an MP_REACH_NLRI containing one or more flow-spec NLRI does not have a valid IPv4 or IPv6 address in its next-hop field, or the

length of the next-hop is 0, then the redirect-to-IP extended community, if present, should be ignored.

The scope of application (in terms of router interfaces/contexts) of the filter rules derived from the redirect-to-IP extended community is outside the scope of this specification except for noting that filter rules derived from VPNv4 and VPNv6 flow-spec routes should only be installed in the VRF contexts that import the routes.

The redirect-to-IP extended community is transitive across AS boundaries. When a flow-spec route with this community is advertised to an EBGP peer the next-hop address in the MP_REACH_NLRI SHOULD be reset to an address of the advertising router by default, per normal BGP procedures. Alternatively, the advertising router MAY be configured to keep the next-hop unchanged, if it is known that the destination AS has a valid route to the next-hop address.

The validation check described in [RFC 5575] and revised in [VALIDATE] SHOULD be applied by default to received flow-spec routes with the redirect-to-IP extended community, as it is to all types of flow-spec routes. This means that a flow-spec route with a destination prefix subcomponent SHOULD NOT be accepted from an EBGP peer unless that peer also advertised the best path for the matching unicast route. BGP speakers that support the redirect-to-IP extended community MUST also, by default, enforce the following check when receiving a flow-spec route from an EBGP peer:

. If the flow-spec route has an IP next-hop X and includes a redirect-to-IP extended community, then the BGP speaker SHOULD discard the redirect-to-ip extended community (and not propagate it further with the flow-spec route) if the last AS in the AS_PATH or AS4_PATH attribute of the longest prefix match for X does not match the AS of the EBGP peer.

It MUST be possible to disable this additional validation check on a per-EBGP session basis.

4. Security Considerations

A system that originates a flow-spec route with a redirect-to-IP extended community can cause many receivers of the flow-spec route to send traffic to a single next-hop, overwhelming that next-hop and resulting in inadvertent or deliberate denial-of-service. This is particularly a concern when the redirect-to-IP extended community is allowed to cross AS boundaries. The validation check described in section 3 significantly reduces this risk.

5. IANA Considerations

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

Type value	Name	Reference
0x0800	Flow spec redirect/mirror to IP next-hop	[this document]

6. References

6.1. Normative References

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

6.2. Informative References

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

[IPV6-FLOW] R. Raszuk, B. Pithawala, D. McPherson, "Dissemination of Flow Specification Rules for IPv6", draft-ietf-idr-flow-spec-v6-00, June 2011.

[VALIDATE] Uttaro, J., Filsfils, C., Mohapatra, P., Smith, D., "Revised Validation Procedure for BGP Flow Specifications", draft-ietf-idr-bgp-flowspec-oid-00, June 2012.

7. Acknowledgments

The authors would like to thank Han Nguyen and Robert Raszuk for their feedback and suggestions.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

James Uttaro
AT&T
200 S. Laurel Avenue
Middletown, NJ 07748
USA
Email: ju1738@att.com

Pradosh Mohapatra
Cisco
170 W. Tasman Drive
San Jose, CA 95134
USA
Email: pmohapat@cisco.com

David Smith
Cisco
111 Wood Avenue South
Iselin, NJ 08830
USA
E-mail: djsmith@cisco.com

Wim Henderickx
Alcatel-Lucent
Copernicuslaan 50
2018 Antwerp, Belgium
Email: wim.henderickx@alcatel-lucent.be

Adam Simpson
Alcatel-Lucent
600 March Road
Ottawa, Ontario K2K 2E6
Canada
Email: adam.simpson@alcatel-lucent.com

Matthieu Texier
Arbor Networks
38 Rue de Berri
75008 Paris
Email: mtexier@arbor.net