

IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: Aug 2, 2015

J. Uttaro
AT&T

J. Haas
Juniper Networks

M. Texier
Arbor Networks

A. Karch
A. Sreekantiah
S. Ray
Cisco Systems

A. Simpson
W. Henderickx
Alcatel-Lucent

Feb 2, 2015

BGP Flow-Spec Redirect to IP Action
draft-ietf-idr-flowspec-redirect-ip-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 Aug 2, 2015.

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 VPN infrastructure.

This draft defines a new redirect-to-IP flow-spec action that provides a simpler method of policy-based forwarding. The details of the action, including the IPv4 or IPv6 target address, are encoded in newly defined BGP extended communities.

Table of Contents

1.	Introduction.....	3
2.	Terminology.....	3
3.	Redirect to IP Extended Communities.....	3
	3.1.	5
	Validation Procedures.....	
4.	Security Considerations.....	6
5.	IANA Considerations.....	6
6.	References.....	6
	6.1.	6
	Normative References.....	
	6.2.	6
	Informative References.....	
7.	Contributors.....	7
8.	Acknowledgments.....	7

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 in one or more BGP extended communities). 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 VPN infrastructure.

This draft proposes a new redirect-to-IP flow-spec action that provides a simpler method of policy-based forwarding. The details of the action, including the IPv4 or IPv6 target address, are encoded in newly defined BGP extended communities.

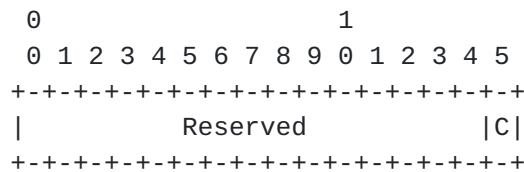
2. Terminology

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 Communities

This document defines two new BGP extended communities. The extended communities have a type indicating they are transitive and IPv4-address-specific or IPv6-address-specific, depending on whether the redirection target address is IPv4 or IPv6. The sub-type value [to be assigned by IANA] indicates that the global administrator and local administrator fields encode a flow-spec 'redirect to IP' action. In the new extended communities the 4-byte or 16-byte global administrator field encodes the IPv4 or IPv6 address that is the redirection target address and the 2-byte local administrator field is formatted as shown in Figure 1.

Figure 1 : Local Administrator



In the local administrator field the least-significant bit is defined as the 'C' (or copy) bit. When the 'C' bit is set the redirection applies to copies of the matching packets and not to the original traffic stream.

All bits other than the 'C' bit in the local administrator field MUST be set to 0 by the originating BGP speaker and ignored by receiving BGP speakers.

When a BGP speaker receives a flow-spec route with a 'redirect to IP' extended community and this route represents the one and only best path, it installs a traffic filtering rule that matches the packets described by the NLRI field and redirects them (C=0) or copies them (C=1) towards the IPv4 or IPv6 address in the extended community's global administrator field (the 'target address'). The BGP speaker is expected to do a longest-prefix-match lookup of the 'target address' in its forwarding information base (FIB) and forward the redirected/copied packets based on the resulting route (the 'target route'). If the 'target route' has multiple ECMP next-hops the redirected/copied packets SHOULD be load-shared across these next-hops according to the router's ECMP configuration. If the 'target route' has one or more tunnel next-hops then the appropriate encapsulations SHOULD be added to the redirected/copied packets. If the 'target address' is invalid or unreachable then the extended community SHOULD be ignored.

If a BGP speaker receives a flow-spec route with multiple 'redirect to IP' extended communities and this route represents the one and only best path, it SHOULD load-share the redirected/copied packets across all the 'target addresses' according to its ECMP configuration. If the BGP speaker is not capable of redirecting and copying the same packet it SHOULD ignore the extended communities with C=0. If the BGP speaker is not capable of redirecting/copying a packet towards multiple 'target addresses' it SHOULD deterministically select one 'target address' and ignore the others.

If a BGP speaker receives multiple flow-spec routes for the same flow-spec NLRI and all of them are considered best and usable paths according to the BGP speaker's multipath configuration and each one carries one or more 'redirect to IP' extended communities, the BGP speaker SHOULD load-share the redirected/copied packets across all the 'target addresses', with the same fallback rules as discussed in the previous paragraph. Note that this situation does not require the BGP speaker to have multiple peers - i.e. Add-Paths could be used for the flow-spec address family.

If a BGP speaker receives a flow-spec route with one or more 'redirect to IP' extended communities and one or more 'redirect to VRF' extended communities, and this route represents the one and only best path, the 'redirect to IP' actions described above should be applied in the context of the 'target VRF' matching the 'redirect to VRF' extended community - i.e. the 'target addresses' should be looked up in the FIB of the 'target VRF'. If there are multiple 'redirect to VRF' extended communities in the route the 'target VRF' SHOULD be the one that matches the 'redirect to VRF' extended community with the highest numerical value. If the BGP speaker is not capable of 'redirect to VRF' followed by 'redirect to IP' then it SHOULD give preference to performing the 'redirect to VRF' action and doing only longest-prefix-match forwarding in the 'target VRF'.

If a BGP speaker receives multiple flow-spec routes for the same flow-spec NLRI and all of them are considered best and usable paths according to the BGP speaker's multipath configuration and they carry a combination of 'redirect to IP' and 'redirect to VRF' extended communities, the BGP speaker SHOULD apply the 'redirect to IP' actions in the context of the 'target VRF' as described above. Note that this situation does not require the BGP speaker to have multiple peers - i.e. Add-Paths could be used for the flow-spec address family.

3.1. Validation Procedures

The validation check described in [[RFC 5575](#)] and revised in [[VALIDATE](#)] SHOULD be applied by default to received flow-spec routes with a '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 extended communities defined in this draft MUST also, by default, enforce the following check when receiving a flow-spec route from an EBGP peer: if the received flow-

spec route has a 'redirect to IP' extended community with a 'target address' X (in the global administrator field) and the best matching route to X is not a BGP route with origin AS matching the peer AS then the extended community should be discarded and not propagated along with the flow-spec route to other peers. 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 3.1](#) significantly reduces this risk.

5. IANA Considerations

This document requests a new sub-type from the "Transitive IPv4-Address-Specific" extended community registry. The sub-type name shall be 'Flow-spec Redirect to IPv4'.

This document requests a new sub-type from the "Transitive IPv6-Address-Specific" extended community registry. The sub-type name shall be 'Flow-spec Redirect to IPv6'.

IANA is requested to deprecate the type 0x0800 type/sub-type.

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. Contributors](#)

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

[8. 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

Jeffrey Haas
Juniper Networks
1194 N. Mathida Ave.
Sunnyvale, CA 94089
USA
Email: jhaas@juniper.net

Andy Karch
Cisco Systems
170 West Tasman Drive
San Jose, CA 95134
USA
Email: akarch@cisco.com

Saikat Ray
Cisco Systems, Inc.
170, West Tasman Drive
San Jose, CA 95134
USA
Email: sairay@cisco.com

Pradosh Mohapatra
Cumulus Networks
Email: pmohapat@cumulusnetworks.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