IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: 6 March 2026

J. Haas
HPE
W. Henderickx
A. Simpson
Nokia

2 September 2025

BGP Flow-Spec Redirect-to-IP Action
draft-ietf-idr-flowspec-redirect-ip-04

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 distributed denial of service (DDoS) mitigation. The flow-spec standard [RFC8955] defines a redirect-to-VRF action for policy-based forwarding. 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

BGP flow-spec [RFC8955] 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 distributed denial of service (DDoS) mitigation.

Every flow-spec route is a rule, consisting of a matching part encoded in the BGP Network Layer Reachability Information (NLRI) field, and an action part encoded in one or more BGP extended communities. Flow-spec defines 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 method for policy-based forwarding to redirect or copy matching traffic toward a specific IP address. This method of redirection and copying is simpler than the existing methods in [RFC8955] and [RFC8956] to redirect traffic to a VRF. The details of the action, including the IPv4 or IPv6 target address, are encoded in newly defined BGP extended communities.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

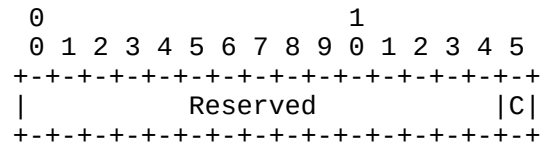
2. Redirect-to-IP Extended Communities

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.

For the IPv4 address-specific extended community [RFC4360], the IANA-assigned sub-type value 0x0c indicates that the Global Administrator and Local Administrator fields encode a flow-spec "redirect-to-IPv4" action. In the encoding of this action, the 4-octet Global Administrator field encodes the IPv4 address that is the redirection target address and the 2-octet local administrator field is formatted as shown in Figure 1.

For the IPv6 address-specific extended community [RFC5701], the IANA-assigned type 0x000c indicates that the Global Administrator and Local Administrator fields encode a flow-spec "redirect-to-IPv6" action. In this encoding, the 16-octet Global Administrator field contains the IPv6 address that is the redirection target address and the 2-octet local administrator field is again 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.

2.1. Validation Procedures

The validation check described in [RFC8955] and [RFC8956], and revised in [RFC9117], 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. When this check is applied, a flow-spec route with a Destination Prefix subcomponent that originated outside the local AS is considered valid only if the neighbor AS implied in the AS_PATH attribute is the neighbor AS of the unicast IP route that is the best match of the destination prefix, and it is also the neighbor AS of all unicast IP routes that are longer matches of the destination prefix.

BGP speakers that support the extended communities defined in this draft MUST also, by default, apply additional validation rules when receiving a flow-spec with these extended communities. More specifically, the router must consider a "redirect-to-IPv4" or "redirect-to-IPv6" extended community to be invalid if the origin AS of the flow-spec route does not match the origin AS of the best-match unicast route for the "target-address". For example:

* If the flow-spec route has a non-empty AS_PATH and any AS_PATH path segment is of the type AS_SET or AS_CONFED_SET, then the extended community is considered "invalid". Compare similarly to Section 4 of [RFC9774].

* If the flow-spec route has a non-empty AS_PATH indicating origin AS = X, and the resolving route of the "target-address" is a BGP route with a non-empty AS_PATH indicating origin AS = X, then the extended community is considered "valid".

* If the flow-spec route has a null/empty AS_PATH, or an AS_PATH with only local confederation elements, and the resolving route of the "target-address" is a BGP route with a null/empty AS_PATH or an AS_PATH with only local confederation elements then the extended community is considered "valid".

* If the flow-spec route has a null/empty AS_PATH, or an AS_PATH with only local confederation elements, and the resolving route of the "target-address" is a non-BGP route, then the extended community is considered "valid".

* If the flow-spec route has a null/empty AS_PATH or an AS_PATH with only local confederation elements, and the resolving route of the "target-address" is a BGP route that originated outside the local AS or confederation, then the extended community is considered "invalid".

* If the flow-spec route has a non-empty AS_PATH indicating origin AS = X, and the resolving route of the "target-address" is a BGP route with a null/empty AS_PATH or an AS_PATH with only local confederation elements, or it is a non-BGP route then the extended community is considered "invalid".

If any of the above checks determine that a "redirect-to-IP" extended community is invalid, the extended community SHOULD be ignored.

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

2.2. Redirecting Matching Flowspec Traffic

Traffic that is to be redirected/copied for a "redirect-to-IP" extended community SHOULD only be redirected if the community type matches the traffic type.

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 the database it uses to resolve next-hop addresses and then 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. (For example, BGP Add-Paths [RFC7911] could be used for the flow-spec address family.)

2.2.1. Interactions with Redirect to VRF Extended Community

If a BGP speaker receives a flow-spec route with the following:

* One or more "redirect-to-IP" extended communities and,

* One or more "redirect-to-VRF" ([RFC8955], Section 7.4) extended communities and,

* This route represents the “one and only” best path,

then 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. BGP Add-Paths [RFC7911] could be used for the flow-spec address family.

2.2.2. Interactions with Other Flowspec Traffic Filtering Actions

Traffic redirection or copying leverages the result of the lookup operation in the database used to resolve next hop addresses of the target address carried in the redirect-to-ip Extended Communities. The forwarding result of this operation typically is implemented as a IP forwarding operation, or results in the matching traffic being encapsulated in a tunnel. This operation will generally short-circuit other traffic filtering options for the redirected or copied traffic. As a result, the expected behaviors when redirect-to-ip is implemented and the following other traffic filtering actions are carried with the flowspec route are:

Traffic Rate in bytes (Section 7.1 of [RFC8955]):
: Redirected and copied traffic are subject to the traffic policing mechanisms resulting from the lookup vs. the next hop database. This traffic filtering action is thus IGNORED for traffic that is redirected or copied.

Traffic Rate in packets (Section 7.2 of [RFC8955]):
: Redirected and copied traffic are subject to the traffic policing mechanisms resulting from the lookup vs. the next hop database. This traffic filtering action is thus IGNORED for traffic that is redirected or copied.

Terminal action (Section 7.3 of [RFC8955]):
: Redirection of matching traffic is considered a terminating action and the non-terminal action (T == 1) is IGNORED. Copying of matching traffic is considered a non-terminating action and the terminal action bit's behavior is respected in implementations that support copying.

Sampling (Section 7.3 of [RFC8955]):
: Sampling MAY be done as part of the redirection/copy.

Traffic marking (Section 7.5 of [RFC8955]):
Redirected and copied traffic are subject to the traffic policing mechanisms resulting from the lookup vs. the next hop database. This traffic filtering action is thus IGNORED for traffic that is redirected or copied.

SFC classifier (Section 7.4 of [RFC8955]):
Redirected and copied traffic are subject to the traffic policing mechanisms resulting from the lookup vs. the next hop database. This traffic filtering action is thus IGNORED for traffic that is redirected or copied.

Redirect to indirection-id (Section 4 of [I-D.ietf-idr-flowspec-path-redirect]):
In general, multiple operations supporting redirection or copying SHOULD NOT be present in the same BGP flowspec route. When both redirect-to-ip and redirect to indirection-id are both present on the same route, redirect-to-IP SHALL have higher precendence for redirection and copying.

SRv6 Redirect to indirection-id (Section 2 of [I-D.ietf0-idr-srv6-flowspec-path-redirect]):
In general, multiple operations supporting redirection or copying SHOULD NOT be present in the same BGP flowspec route. When both redirect-to-ip and SRv6 redirect to indirection-id are both present on the same route, redirect-to-IP SHALL have higher precendence for redirection and copying.

3. 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 Section 2.1 significantly reduces this risk.

4. Implementation Status

This section documentations the [RFC7942] implementation status of this document.

4.1. HPE / Juniper Networks

Organization:
HPE / Juniper Networks

Implementation Name:
Junos 18.4R1 and later

Description:
Juniper redirect-to-ip feature

Maturity:
Widely used.

Coverage:
* Section 2 IPv4 Extended Community - Implemented.

* Section 2 IPv6 Extended Community - Not Implemented.

* Section 2 Redirect (C == 0) - Implemented.

* Section 2 Copy (C == l) - Not Implemented.

* Section 2.1 Validation - Not Implemented.

* Section 2.2 Longest prefix match - Implemented.

* Section 2.2 Best path ECMP - Implemented.

* Section 2.2 Multiple communities ECMP load sharing - Implemented.

* Section 2.2 Redirect-to-IP in Redirect-to-VRF - Not Implemented.

Version Compatibility:
draft-ietf-idr-flowspec-redirect-ip-02

Licensing:
Proprietary

Implementation Experience:

Contact Information:
Jeffrey Haas - jhaas@juniper.net

Last Updated:
August 2025

4.2. Cisco IOS-XR

Organization:
Cisco

Implementation Name:
Cisco IOS-XR

Description:

Maturity:
Widely used.

Coverage:
* Section 2 IPv4 Extended Community - Not Implemented.

* Section 2 IPv6 Extended Community - Not Implemented.

* Section 2 Redirect (C == 0) - Not Implemented.

* Section 2 Copy (C == l) - Not Implemented.

* Section 2.1 Validation - Not Implemented.

* Section 2.2 Longest prefix match - ?.

* Section 2.2 Best path ECMP - ?.

* Section 2.2 Multiple communities ECMP load sharing - Not Implemented.

* Section 2.2 Redirect-to-IP in Redirect-to-VRF - Not Implemented.

Version Compatibility:
draft-ietf-idr-flowspec-redirect-ip-00

Licensing:
Proprietary

Implementation Experience:

Contact Information:
Jakob Heitz

Last Updated:
14 October 2024 - IDR mailing list

5. IANA Considerations

IANA has allocated an extended community from the "Transitive IPv4-Address-Specific Extended Community Sub-Types" registry. The Sub-Type value is 0x0c. The Name shall be "Flow-spec Redirect-to-IPv4". The Reference shall be this document.

IANA has allocated an extended community from the "Transitive IPv6-Address-Specific Extended Community Types" registry. The Type value is 0x000c. The Name shall be "Flow-spec Redirect-to-IPv6". The Reference shall be this document.

In a previous draft of this document, IANA had allocated an extended community from the "BGP Transitive Extended Community Types" registry with Type Value 0x08 for "Flow spec redirect/mirror to IP next-hop". IANA is requested to deprecate this registration.

6. References

6.1. 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>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <https://www.rfc-editor.org/info/rfc4360>.

[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>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8955] Loibl, C., Hares, S., Raszuk, R., McPherson, D., and M. Bacher, "Dissemination of Flow Specification Rules", RFC 8955, DOI 10.17487/RFC8955, December 2020, <https://www.rfc-editor.org/info/rfc8955>.

[RFC8956] Loibl, C., Ed., Raszuk, R., Ed., and S. Hares, Ed., "Dissemination of Flow Specification Rules for IPv6", RFC 8956, DOI 10.17487/RFC8956, December 2020, <https://www.rfc-editor.org/info/rfc8956>.

[RFC9117] Uttaro, J., Alcaide, J., Filsfils, C., Smith, D., and P. Mohapatra, "Revised Validation Procedure for BGP Flow Specifications", RFC 9117, DOI 10.17487/RFC9117, August 2021, <https://www.rfc-editor.org/info/rfc9117>.

[RFC9774] Kumari, W., Sriram, K., Hannachi, L., and J. Haas, "Deprecation of AS_SET and AS_CONFED_SET in BGP", RFC 9774, DOI 10.17487/RFC9774, May 2025, <https://www.rfc-editor.org/info/rfc9774>.

[I-D.ietf-idr-flowspec-path-redirect] Van de Velde, G., Patel, K., and Z. Li, "Flowspec Indirection-id Redirect", Work in Progress, Internet-Draft, draft-ietf-idr-flowspec-path-redirect-12, 24 November 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-flowspec-path-redirect-12>.

[I-D.ietf0-idr-srv6-flowspec-path-redirect] Van de Velde, G., Patel, K., Li, Z., and H. Chen, "Flowspec Indirection-id Redirect for SRv6", Work in Progress, Internet-Draft, draft-ietf0-idr-srv6-flowspec-path-redirect-12, 21 October 2024, <https://datatracker.ietf.org/doc/html/draft-ietf0-idr-srv6-flowspec-path-redirect-12>.

6.2. Informative References

[RFC7911] Walton, D., Retana, A., Chen, E., and J. Scudder, "Advertisement of Multiple Paths in BGP", RFC 7911, DOI 10.17487/RFC7911, July 2016, <https://www.rfc-editor.org/info/rfc7911>.

[RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <https://www.rfc-editor.org/info/rfc7942>.

Acknowledgements

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

Contributors

James Uttaro
Individual Contributor
Email: juttaro@ieee.org

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

Saikat Ray
Individual Contributor
Email: raysaikat@gmail.com

David Smith
Cisco
111 Wood Avenue South
Iselin, NJ 08830
United States of America
Email: djsmith@cisco.com

Pradosh Mohapatra
Individual Contributor
Email: pradosh@google.com

Matthieu Texier
Pragma Security
Email: matthieu@pragma-security.com

Authors' Addresses

Jeffrey Haas
HPE
1133 Innovation Way
Sunnyvale, CA 94089
United States of America
Email: jhaas@juniper.net

Wim Henderickx
Nokia
copernicuslaan 50
2018 Antwerp
Belgium
Email: wim.henderickx@nokia.com

A. Simpson
Nokia
Email: adam.1.simpson@nokia.com