Workgroup: IDR Working Group
Internet-Draft:
draft-ietf-idr-flowspec-path-redirect-12
Published: 24 November 2022
Intended Status: Standards Track
Expires: 28 May 2023
Authors: G. Van de Velde, Ed. K. Patel Z. Li
Nokia Arrcus Huawei Technologies

Flowspec Indirection-id Redirect

Abstract

This document defines a new extended community known as "FlowSpec Redirect to indirection-id Extended Community". This extended community triggers advanced redirection capabilities to flowspec clients. When activated, this flowspec extended community is used by a flowspec client to retrieve the corresponding next-hop and encoding information within a localised indirection-id mapping table.

The functionality detailed in this document allows a network controller to decouple the BGP flowspec redirection instruction from the operation of the available paths.

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 28 May 2023.

Copyright Notice

Table of Contents

1. Introduction

Flowspec 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 flowspec standard [RFC8955] defines a redirect-to-VRF action for policy-based forwarding, but this mechanism is not always sufficient, particularly if the redirected traffic needs to be steered onto an explicit path.

Every flowspec policy route is effectively a rule, consisting of two parts. The first part, encoded in the NLRI field, provides information about the traffic matching the policy rule. the second part, encoded in one or more BGP extended communities, provides policy instructions for traffic handling on the flowspec client. The flowspec standard [RFC8955] 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 to steer traffic towards an alternate destination is useful for DDoS mitigation, however using this methodology can be cumbersome when

there is need to steer the traffic onto an explicitely defined traffic path.

This draft specifies a "Redirect to indirection-id" flowspec action making use of a 32-bit indirection-id using a new extended community. Each indirection-id serves as anchor point, for policy-based forwarding onto an explicit path by a flowspec client.

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. indirection-id and indirection-id table

The indirection-id is a 32-bit unsigned number, used as anchor point on a flowspec client for policy-based forwarding onto an explicit path by a flowspec client.

The indirection-id table is the table construct of indirection-id values, grouped by indirection-id "ID-Type". Each entry in this table contains policy-based forwarding and encoding instructions.

The configuration of the indirection-id table on a flowspec client is a localised operation on each router, and MAY happen out-of-band from BGP flowspec. For some use-case scenarios the indirection-id "ID-Type" provides additional (maybe even fully sufficient) context for a flowspec client for policy based forwarding, making a localised indirection-id table obsolete. For example, when the indirection-id refers to a MPLS segment routing node-id [RFC9256], then the indirection-id provides sufficient information for a segment routing lookup on the flowspec client.

3. Use Case Scenarios

This section describes a few use-case scenarios when deploying "Redirect to indirection-id".

3.1. Redirection shortest Path tunnel

Description:

The first use-case describes an example where a single flowspec route is sent from a BGP flowspec controller to many BGP flowspec clients. This BGP flowspec route carries the "Redirect to indirection-id" to all flowspec clients with intent to redirect

matching dataflows onto a shortest-path tunnel pointing towards a single remote destination.

In this first use-case scenario, each flowspec client receives flowspec routes. The received flowspec routes have the extended "Redirect to indirection-id" community attached. Each "Redirect to indirection-id" community embeds two relevant components: (1) 32-bit indirection-id and (2) ID-type. These two components provide the flowspec client with sufficient information for policy based forwarding, with intent to steer and encapsulate the data-packet accordingly upon a shortest path tunnel to a single remote end-point.

Requirements:

For redirect to shortest path tunnel it is required that the tunnel MUST be operational and allow packets to flow between tunnel head- and tail-end.

Example: Indirection-ID community "ID-Type" which can be used:

* 0 (localised ID): When the intent is to use a localised Indirection-id table, configured through out-of-band procedures.

* 1 or 2 (Node ID's): This type can be used when the goal is to use MPLS based Segment Routing towards a remote destination. In this use-case scenario the flowspec rule contains a SR (Segment Routing) node SID to steer traffic towards.

3.2. Redirection to path-engineered tunnels

Description:

The second use-case describes an example where a single flowspec route is sent from a BGP flowspec controller to many BGP flowspec clients. This BGP flowspec route carries policy information to steer traffic upon a path-engineered tunnel. It is assumed that the path engineered tunnels are configured using out-of-band from BGP flowspec.

Segment Routing Example:

For this example the indirection-id "ID-Type" points towards a Segment Routing Binding SID. The Binding SID is a segment identifier value (as per segment routing definitions in [RFC8402]) used to associate an explicit path. The Binding SID and the associated path engineered tunnel may for example be setup by a controller using BGP as specified in [SR-TE] or alternately by using PCEP as detailed in [RFC8664]. To conclude, when a BGP speaker at some point in time receives a flowspec route with an extended "Redirect to indirection-

id' community, it installs a policy-based forwarding rule to redirect packets onto an explicit path, associated with the corresponding Binding SID. The encoding of the Binding SID within the "Redirect to indirection-id" extended community is specified in section 4.

Requirements:

For redirect to path engineered tunnels it is required that the tunnel MUST be operational and allow packets to flow over the engineered path between tunnel head- and tail-end.

Example: Indirection-ID community "ID-Type" to be used:

- 0 (localised ID): When the intent is to policy-based steer traffic using Indirection. The engineered path is configured through out-of-band procedures and uses the 32-bit Indirection-id as local anchor point on the local flowspec client.

- 3 or 4 (Binding Segment ID's): This type can be used when the goal is to use MPLS based Segment Routing towards an out-of-band configured explicit path.

- 5 (Tunnel ID): When the intent is to policy-based steer traffic using a global tunnel-id. The engineered path is configured through out-of-band procedures and uses the 32-bit Indirection-id as global anchor point on the local flowspec client.

3.3. Redirection to complex dynamically constructed tunnels

Description:

A third use-case describes the application and redirection towards complex dynamically constructed tunnels. For this use-case a BGP flowspec controller injects a single flowspec route with two unique "Redirect to indirection-id" communities attached, each community tagged with a different Sequence-ID (S-ID). A flowspec client should use the Sequence-ID (S-ID) to sequence the received flowspec redirect information. A potential use-case scenario would for example be the dynamic construction of Segment Routing Central Egress Path Engineered tunnel [RFC9087] or next-next-hop tunnels.

Segment Routing Example:

i.e. a classic Segment Routing example using complex tunnels is found in DDoS mitigation and traffic offload. Suspicious traffic (e.g. dirty traffic flows) may be policy-based routed into a purpose built Segment Routing Central Egress Path Engineered tunnel [RFC9087]. For this complex dynamic redirect tunnel construct, a first "Redirect to indirection-id" (i.e. S-ID=0) may be used to

redirect traffic into a tunnel towards a particular egress router, while a second "Redirect to indirection-id" (i.e. S-ID=1) is used to steer traffic beyond the particular egress router towards a pre-identified interface/peer. From data-plane perspective, the principles documented by [RFC9087] are valid for this use case scenario.

Requirements:

To achieve redirection towards complex dynamically constructed tunnels, multiple "Redirect to indirection-id" communities are imposed upon the flowspec route. The "Redirect to indirection-id" communities should be sequenced using the Sequence ID (S-ID). For redirect to complex dynamic engineered tunnels the tunnel MUST be operational and allow packets to flow over the engineered path between tunnel head- and tail-end.

Example: Indirection-ID community "ID-Type" to be used:

* 0 (localised ID) with S-ID: When the intent is to construct a dynamic engineered tunnel, then a sequence of localised indirection-ids may be used. The Sequence ID (S-ID) MUST be used to sequence multiple "Redirect to indirection-id" actions to construct a more complex engineered tunnel. The creation of the localised indirection-id table is operationalised out-of-band and is outside scope of this document.

4. Redirect to indirection-id Community

This document defines a new transitive BGP extended community known as "FlowSpec Redirect to indirection-id Extended Community" with the Type and the Sub-Type field to be assigned by IANA. The format of this extended community is show in Figure 1.

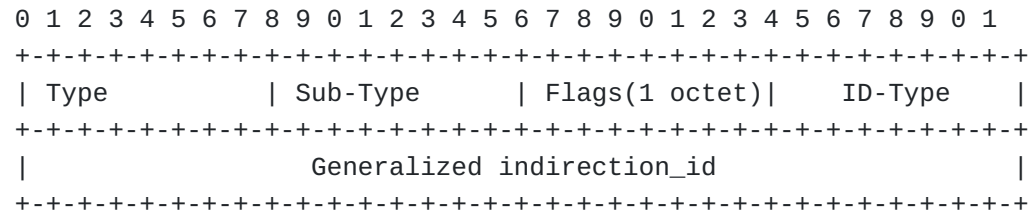

Figure 1: Extended-Community

The meaning of the extended community fields are as follows:

* Type: 1 octet to be assigned by IANA.

*Sub-Type: 1 octet to be assigned by IANA.

*Flags: 1 octet field. Following Flags are defined.

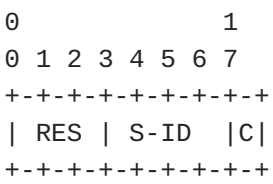

Figure 2: Flags

The least-significant Flag 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.

The 'S-ID' field identifies a 4 bit Sequence ID field. This field is used to provide a flowspec client an indication how and where to sequence the received indirection-ids. The Sequence ID value 0 indicates that Sequence ID field is NOT set and all other sequence ID's SHOULD be ignored. A single flowspec rule MUST NOT have more as one indirection-id per S-ID. On a flowspec client the indirection-id with lowest S-ID MUST be imposed first for any given flowspec entry.

All bits other than the 'C' and 'S-ID' bits MUST be set to 0 by the originating BGP speaker and ignored by receiving BGP speakers.

ID-Type: 1 octet value. This draft defines following Context Types:

*0 - Localised ID (The flowspec client uses the received 32-bit indirection-id to lookup forwarding information within the localised indirection-id table. The allocation and programming of the localised indirection-id table is outside scope of the document)

*1 - Node ID with SID/index in MPLS-based Segment Routing (This means the 32-bit indirection-id is mapped to an MPLS label using the index as a global offset in the SID/label space)

*2 - Node ID with SID/label in MPLS-based Segment Routing (This means the 32-bit indirection-id is mapped to an MPLS label using the 32-bit indirection-id as global label)

*3 - Binding Segment ID with SID/index in MPLS-based Segment Routing (This means the 32-bit indirection-id is mapped to an

MPLS binding label using the indirection-id as index for global offset in the SID/label space) [RFC8402]

*4 - Binding Segment ID with SID/label in MPLS-based Segment Routing (This means 32-bit indirection-id is mapped to an MPLS binding label using the 32-bit indirection-id as global label) [RFC8402]

*5 - Tunnel ID (Tunnel ID is within a single administrative domain a 32-bit globally unique tunnel identifier. The allocation and programming of the Tunnel ID within the localised indirection-id table is outside scope of the document)

Generalized indirection_id: 32-bit identifier used as indirection_id

5. Redirect using localised indirection-id mapping table

When a BGP flowspec client receives a flowspec policy route with a "Redirect to indirection-id" extended community attached, and the route represents the best BGP path, it will install a flowspec policy-based forwarding rule matching the tupples described by the flowpsec NLRI field and consequently redirects the flow (C=0) or copies the flow (C=1) using the information identified by the "Redirect to indirection-id" community.

6. Validation Procedures

The validation check described in [RFC8955] SHOULD be applied by default by a flowspec client, for received flowspec policy routes containing a "Redirect to indirection-id" extended community. This results that a flowspec 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.

While it MUST NOT happen, and is seen as invalid combination, it is possible from a semantics perspective to have multiple clashing redirect actions defined within a single flowspec rule. For best and consistant compatibility with legacy implementations, the redirect functionality as documented by rfc8955 MUST NOT be broken, and hence when a clash occurs, then rfc8955 based redirect MUST take priority. Additionally, if the "Redirect to indirection-id" does not result in a valid redirection, then the flowspec rule MUST be processed as if the "Redirect to indirection-id" community was not attached to the flowspec route. In addition the flowspec client MUST provide an indication that the respective "'Redirect to indirection-id" resulted in an invalid redirection action.

7. IANA Considerations

This document requests a new Transitive Extended Community Type and a new registery sub-type. The new Transitive Extended Community Type name shall be "FlowSpec Redirect to indirection-id Extended Community (Sub-Types are defined in the "FlowSpec Redirect to indirection-id Extended Community Sub-Type" registery)". The name of the new Sub-type registery shall be "FlowSpec Redirect to indirection-id Extended Community Sub-Type"

Under "Transitive Extended Community:"

Registry: "FlowSpec Redirect to indirection-id Extended Community (Sub-Types are defined in the "FlowSpec Redirect to indirection-id Extended Community Sub-Type" registery)"

Registration Procedure(s): First Come, First Served

0x09 FlowSpec Redirect to indirection-id Extended Community

New Sub-Type Registry: "FlowSpec Redirect to indirection-id Extended Community Sub-Type"

Value	Code	Reference
0x00	Flowspec Redirect to 32-bit Path-id	RFC-To-Be

Figure 3: IANA-Flags

8. Security Considerations

A system using "Redirect to indirection-id" extended community can cause during the redirect mitigation of a DDoS attack overflow of traffic received by the mitigation infrastructure.

9. References

9.1. 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>.

[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>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[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>.

9.2. Informative References

[RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <https://www.rfc-editor.org/info/rfc8664>.

[RFC9087] Filsfils, C., Ed., Previdi, S., Dawra, G., Ed., Aries, E., and D. Afanasiev, "Segment Routing Centralized BGP Egress Peer Engineering", RFC 9087, DOI 10.17487/RFC9087, August 2021, <https://www.rfc-editor.org/info/rfc9087>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

[SR-TE] Talaulikar, K., Filsfils, C., Previdi, S., Jain, D., Mattes, P., and S. Lin, "Advertising Segment Routing Policies in BGP", July 2022.

Acknowledgements

This document received valuable comments and input from IDR working group including Adam Simpson, Mustapha Aissaoui, Jan Mertens, Robert Raszuk, Jeff Haas, Susan Hares and Lucy Yong.

Contributors

The following people contributed to the content of this document and should be considered as co-authors:

*Arjun Sreekantiah, USA, arjunhrs@gmail.com

*Nan Wu, Huawei Technologies, China, eric.wu@huawei.com

*Shunwan Zhuang, Huawei Technologies, China, zhuangshunwan@huawei.com

*Wim Henderickx, Nokia Antwerp, Belgium, wim.henderickx@nokia.com

Authors' Addresses

Gunter Van de Velde (editor)
Nokia
Antwerp
Belgium

Email: gunter.van_de_velde@nokia.com

Keyur Patel
Arrcus
United States of America

Email: keyur@arrcus.com

Zhenbin Li
Huawei Technologies
Huawei Bld., No. 156 Beiquing Rd
Beijing
100095
China

Email: lizhenbin@huawei.com