IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 3, 2018

G. Van de Velde, Ed.
Nokia
K. Patel
Arrcus
Z. Li
Huawei Technologies
August 30, 2017

Flowspec Indirection-id Redirect
draft-ietf-idr-flowspec-path-redirect-02

Abstract

This document defines a new extended community known as flowspec redirect-to-indirection-id. This extended community triggers advanced redirection capabilities to flowspec clients. When activated, this flowspec extended community is used by a flowspec client to find the correct next-hop 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 actual redirection path selected.

Requirements Language

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 [1].

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 http://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 March 3, 2018.

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 flow-spec standard RFC5575 [2] 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 explicite path.

Every flowspec policy 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 RFC5575 [2] 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 but using this technology can be cumbersome when there is need to steer the traffic onto an explicitely defined traffic path.

This draft proposes a new redirect-to-indirection-id flowspec action making use of a 32-bit indirection-id within a new extended community. Each indirection-id serves as anchor point, for policy-based forwarding onto an explicite path on a flowspec client.

A flowspec based indirection service plane can be create when a single 32-bit flowspec indirection-id maps towards a pool of explicite paths.

2. indirection-id and indirection-id table

The indirection-id is a 32-bit unsigned number, used as anchor point on a flowspec client. The indirection-id is on a flowspec client the lookup key-value within a localised list of potential indirection paths. The indirection-id will allow the flowspec client to steer traffic to a particular path or into an indirection service plane by doing a recursive key-value lookup.

The indirection-id table is the table containing an ordered list of indirection-id key-values, ordered by indirection-id type; where each key-value maps towards a particular path or set of paths. The indirection-id type MAY provide additional context about the indirection-id 32-bit value. The flowspec client MUST use the indirection-id as key-value within the indirection-id type corresponding indirection-id table to locate the explicite path and corresponding next-hop information.

The configuration of the indirection-id table on a flowspec client MAY happen out-of-band from BGP flowspec and is a localised construct on each router. For some use-case scenarios the indirection-id type provides additional (maybe even fully sufficient) context towards a flowspec client to deduct automatic, without explicite out-of-band configuration, the indirection-id table. For example, when the indirection-id refers to a segment routing node-id [6], then indirection-id type can provide the flowspec client the awareness that the indirection-id is a segment routing node-id. For this example the indirection-id type allows the flowspec clients to do a recursive lookup using traditional segment routing technology.

To summarise, each indirection-id key-value entry in the indirection-table maps recursively to sufficient next-hop information (parameters regarding encapsulation, egress-interface, QoS, etc...) to successfully indirect traffic according flowspec controller expectations.

3. Use Case Scenarios

This section describes 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 to redirect matching dataflows onto a shortest-path tunnel pointing towards a single remote destination.

For this first use-case scenario, each flowspec client receives flowspec routes. The 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 key-value and (2) indirection-id type. The indirection-id type is used to identify the corresponding indirection-id table, and the actual 32-bit indirection-id key-value is used within the indirection-id table to locate the corresponding next-hop information. The finite result of this operation is sufficient tunnel encapsulation information to forward and encapsulate the data-packet accordingly to a remote tunnel end-point.

Requirements:

For redirect to shortest path tunnel it is required that the tunnel MUST be up-and-running and allow packets to be unidirectional exchanged between tunnel head- and tail-end.

Example: Indirection-ID community types to be used:

- 0 (localised ID): When the intent is to use a localised Indirection-id table on the flowspec client. This requires out-of-band configuration of the indirection-id table

- 1 (Node ID): When the intent is to use a Segment Routing based Indirection-id table on the flowspec client. This requires that Segment Routing is enabled on the flowspec client.

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 the redirect-to-indirection-id extended community to all flowspec clients with instructions to redirect matching dataflows onto a path engineered tunnel. It is expected that each of the path engineered tunnels is instantiated by out-of-band configuration and can be uniquely identified by the combination of (1) indirection-id 32-bit key-value and (2) indirection-id type.

For this second use-case scenario, each flowspec client receives flowspec routes. The flowspec routes have the extended redirect-to-indirection-id community attached. Each redirect-to-indirection-id community embeds two relevant components similar as explained in previous use-case. However the finite result of this operation is sufficient tunnel encapsulation information to forward and encapsulate the data-packet accordingly to a remote tunnel end-point using a path engineered tunnel construction.

Segment Routing Example:

For this example the indirection-id type informs the flowspec client that the indirection-id 32-bit key-value references a Segment Routing Binding SID. The Binding SID is a segment identifier value (as per segment routing definitions in [I-D.draft-ietf-spring-segment-routing] [6]) used to associate an explicit path. The Binding SID and corresponding path engineered tunnel can for example be setup by a controller using BGP as specified in [I-D.sreekantiah-idr-segment-routing-te] [5] or by using PCEP as detailed in draft-ietf-pce-segment-routing [7]. To conclude, when a BGP speaker at some point in time receives a flow-spec route with an extended 'redirect-to-indirection-id' community, it installs a traffic filtering rule that matches particular packets and redirects them 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 engineered tunnel MUST be active and allow packets to be unidirectional exchanged between tunnel head- and tail-end.

Example: Indirection-ID community types to be used:

- o 0 (localised ID): When the intent is to use a localised Indirection-id table on the flowspec client. This requires out-of-band configuration of the indirection-id table.

- o 6 (Binding Segment ID): When the intent is to use a Segment Routing based Indirection-id table on the flowspec client. This requires out-of-band configuration of the Binding Segment IDs.

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 flowspec route with two 'redirect-to-indirection-id' communities attached, each tagged with a different Table-ID (TID). A flowspec client may use the Table-ID (TID) to sequence the flowspec redirect information. A common use-case scenario would for example be the dynamic construction of Segment Routing Central Egress Path Engineered tunnel [4] 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 steered into a Segment Routing Central Egress Path Engineered tunnel [4]. For this complex dynamic redirect tunnel construction, a first redirect-to-indirection-id (i.e. TID=0) is used to redirect traffic into a tunnel towards a particlar egress router, while a second redirect-to-indirection-id (i.e. TID=1) is used to steer traffic beyond the particular egress router towards a pre-identified interface/peer.

For this DDoS use-case, in its simplest embodiment, the flowspec client must dynamically append 2 MPLS Segment Routing labels. A first MPLS Segment Routing label (the outer label) to steer the packet to the egress node (and hence use a shortest path tunnel), while a second MPLS label (matching redirect-to-indirection-id with TID=1), the inner label, to steer on the egress router the original packet to a pre-defined interface/peer. The basic data-plane principles are documented by [4].

Requirements:

To achieve redirection towards complex dynamically constructed tunnels, for each flowspec route, multiple indirection-ids, each using a unique Tunnel ID are pushed upon a given flowspec policy

rule. It is required that there is synchronisation established between the data-plane and control-plane of all relevant devices involved. Each complex dynamically constructed tunnel MUST be operational and allow packets to be unidirectional exchanged between tunnel head- and tail-end before it can be used to redirect traffic.

Example: Indirection-ID community types to be used:

- 0 (localised ID) with TID: When the intent is to use a localised Indirection-id table, then the TID (Table-ID) MUST be used to sequence multiple redirect-to-indirection-id actions to construct a more complex path engineered tunnel. The order of sequencing the redirection information MUST be identified by using the TID field.

4. Redirect to indirection-id Community

This document defines a new BGP extended community known as a Redirect-to-indirection-id extended community. This extended community is a new transitive 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.

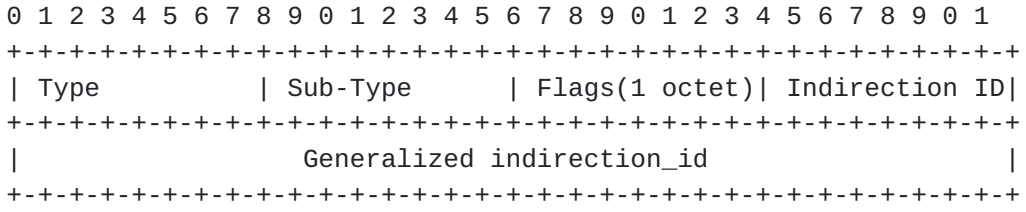

Figure 1

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.

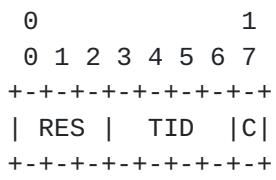

Figure 2

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 'TID' field identifies a 4 bit Table-id field. This field is used to provide the flowspec client an indication how and where to sequence the received indirection-ids to redirecting traffic. TID value 0 indicates that Table-id field is NOT set and SHOULD be ignored. On a flowspec client the indirection-id with lowest TID MUST be processed first for a flowspec route.

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

Indirection ID: 1 octet value. This draft defines following indirection_id Types:

0 - Localised ID (The flowspec client uses the received indirection-id to lookup the redirection information in the localised indirection-id table.)

1 - Node ID (The flowspec client uses the received indirection-id as a Segment Routing Node ID to redirect traffic towards)

6 - Binding Segment ID (The flowspec client uses the received indirection-id as a Segment Routing Binding Segment ID to redirect traffic towards) [I-D.draft-ietf-spring-segment-routing] [6]

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 traffic filtering rule matching the IP 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 RFC5575 [2] and revised in [3] SHOULD be applied by default to received flow-spec routes with a 'redirect to indirection-id' extended community. 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.

While it MUST NOT happen, and is seen as invallid combination, it is possible from a semenatics perspective to have multiple clashing redirect actions defined within a single flowspec rule. For best and consistant RFC5575 flowspec redirect behavior the redirect as documented by RFC5575 MUST not be broken, and hence when a clash occurs, then RFC5575 based redirect SHOULD 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 and MUST provide an indication within the BGP routing table that the respective 'redirect to indirection-id' resulted in an invalid redirection action.

7. Security Considerations

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

8. 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.

9. Contributor Addresses

Below is a list of other contributing authors in alphabetical order:

Arjun Sreekantiah
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: asreekan@cisco.com

Nan Wu
Huawei Technologies
Huawei Bld., No. 156 Beiquing Rd
Beijing 100095
China

Email: eric.wu@huawei.com

Shunwan Zhuang
Huawei Technologies
Huawei Bld., No. 156 Beiquing Rd
Beijing 100095
China

Email: zhuangshunwan@huawei.com

Wim Henderickx
Nokia
Antwerp
BE

Email: wim.henderickx@nokia.com

Figure 3

10. IANA Considerations

This document requests a new type and sub-type for the Redirect to indirection-id Extended community from the "Transitive Extended community" registry. The Type name shall be "Redirect to indirection-id Extended Community" and the Sub-type name shall be 'Flow-spec Redirect to 32-bit Path-id'.

In addition, this document requests IANA to create a new registry for Redirect to indirection-id Extended Community INDIRECTION-IDs as follows:

Under "Transitive Extended Community:"

Registry: "Redirect Extended Community indirection_id"

Reference: [RFC-To-Be]

Registration Procedure(s): First Come, First Served

Registry: "Redirect Extended Community indirection_id"

Value	Code	Reference
0	Localised ID	[RFC-To-Be]
1	Node ID	[RFC-To-Be]
6	Tunnel ID (Tunnel Binding ID)	[RFC-To-Be]

Figure 4

11. References

11.1. Normative References

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997, <http://xml.resource.org/public/rfc/html/rfc2119.html>.

[2] 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>.

11.2. Informative References

[3] Uttaro, J., Filsfils, C., Alcaide, J., and P. Mohapatra, "Revised Validation Procedure for BGP Flow Specifications", January 2014.

[4] Filsfils, C., Previdi, S., Aries, E., Ginsburg, D., and D. Afanasiev, "Segment Routing Centralized Egress Peer Engineering", October 2015.

[5] Sreekantiah, A., Filsfils, C., Previdi, S., Sivabalan, S., Mattes, P., and S. Lin, "Segment Routing Traffic Engineering Policy using BGP", October 2015.

[6] Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., Shakir, R., Bashandy, A., Horneffer, M., Henderickx, W., Tantsura, J., Crabbe, E., Milojevic, I., and S. Ytti, "Segment Routing Architecture", December 2015.

[7] Sivabalan, S., Medved, M., Filsfils, C., Litkowski, S., Raszuk, R., Bashandy, A., Lopez, V., Tantsura, J., Henderickx, W., Hardwick, J., Milojevic, I., and S. Ytti, "PCEP Extensions for Segment Routing", December 2015.

Authors' Addresses

Gunter Van de Velde (editor)
Nokia
Antwerp
BE

Email: gunter.van_de_velde@nokia.com

Keyur Patel
Arrcus
USA

Email: keyur@arrcus.com

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

Email: lizhenbin@huawei.com