IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 3, 2017

G. Van de Velde, Ed.
Nokia
K. Patel
Cisco Systems
Z. Li
Huawei Technologies

August 30, 2016

Flowspec Indirection-id Redirect
draft-ietf-idr-flowspec-path-redirect-00

Abstract

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 into an engineered path or into a service plane.

This document defines a new extended community known as redirect-to-indirection-id (32-bit) flowspec action to provide advanced redirection capabilities on flowspec clients. When activated, the flowspec extended community is used by a flowspec client to find the correct next-hop entry within a localised indirection-id mapping table.

The functionality present in this draft allows a network controller to decouple flowspec functionality from the creation and maintainance of the network's service plane itself including the setup of tunnels and other service constructs that could be managed by other network devices.

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.

Copyright Notice

Table of Contents

1. Introduction

Flowspec RFC5575 [2] is an extension to BGP that allows for the dissemination of traffic flow specification rules. This has many possible applications, however the primary one for many network operators is the distribution of traffic filtering actions for DDoS mitigation.

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 engineered traffic path.

This draft proposes a new redirect-to-indirection-id flowspec action facilitating an anchor point for policy-based forwarding onto an engineered path or into a service plane. The flowspec client consuming and utilizing the new flowspec indirection-id extended-community finds the redirection information within a localised indirection-id mapping table. The localised mapping table is a table construct providing at one side the table key and at the other side next-hop information. The table key consists out the combination of indirection-id type and indirection-id 32-bit value.

The redirect-to-indirection-id flowspec action is encoded in a newly defined BGP extended community. In addition, the type of redirection can be configured as an extended community indirection-id type field.

This draft defines the indirection-id extended-community and the wellknown indirection-id types. The specific solution to construct the localised indirection-id mapping table are out-of-scope of this document.

2. indirection-id and indirection-id table

An indirection-id is an abstract number (32-bit value) used as identifier for a localised indirection decision. The indirection-id will allow a flowspec client to redirect traffic into a service plane or onto an engineered traffic path. For example, when a BGP flowspec controller signals a flowspec client the indirection-id extended community, then the flowspec client uses the indirection-id to make a recursive lookup to retrieve next-hop information found in a localised indirection mapping table.

The indirection-id table is a router localised table. The indirection-id table is constructed out of table keys mapped to flowspec client localised redirection information. The table key is created by the combination of the indirection-id type and the indirection-id 32-bit value. Each entry in the indirection-table key maps to sufficient information (parameters regarding encapsulation, interface, QoS, etc...) to successfully redirect traffic.

3. Use Case Scenarios

This section describes use-case scenarios when deploying redirect-to-indirection-id.

3.1. Redirection shortest Path tunnel

Possible Indirection-ID type examples:

- When deploying on flowspec client a localised Indirection-id table: 0 (localised ID)
- When deploying on flowspec client a Segment Routing based Indirection-id table: 1 (Node ID)

Description:

A first use-case is allowing a BGP Flowspec controller to send a single flowspec policy route (i.e. flowspec_route#1) to many BGP flowspec clients. This flowspec route signals the Flowspec clients to redirect traffic onto a tunnel towards a single IP destination address.

For this first use-case scenario, the flowspec client receives from the flowspec controller a flowspec route (i.e. flowspec_route#1) including the redirect-to-indirection-id extended community. The redirect-to-indirection-id extended community contains the key (indirection-id type + indirection-id 32-bit value) to select the corresponding next-hop information from the flowpsec client localised indirection-id table. The resulting next-hop information for this use-case is a remote tunnel end-point IP address with accordingly sufficient tunnel encapsulation information to forward the packet accordingly.

Requirements:

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

3.2. Redirection to path-engineered tunnels

Possible Indirection-ID type examples:

- When deploying on flowspec client a localised Indirection-id table: 0 (localised ID)

- When deploying on flowspec client a Segment Routing based Indirection-id table: 6 (Binding Segment ID)

Description:

For a second use-case, it is expected that the flowspec client redirect traffic matches the flowspec rule, onto a path engineered tunnel. The path engineered tunnel on the flowspec client SHOULD be created by out-of-band mechanisms. Each path engineered tunnel deployed for flowspec redirection, has a unque key as an identifier. consequently, the key (=indirection-id type and indirection-id 32-bit value) uniquely identifies a single path engineered tunnel on the flowspec client. The localised indirection-id mapping table is the collection of all keys corresponding all path engineered tunnels on the flowspec client.

For this second use-case scenario, the flowspec controller sends a flowspec route (i.e. flowspec_route#2) to some flowspec clients. The flowspec clients, respectively receive the flowspec route. The redirect-to-indirection-id extended community contains sufficient information for the flowspec clients to select the corresponding path-engineered tunnel and consequently provides sufficient tunnel encapsulation information to redirect the packet according the flowspec controller expectations.

Segment Routing Example:

A concrete Segment Routing use-case example of this use-case can be found in segment routed networks where path engineered tunnels can be setup by means of a controller signaling explicit paths to peering routers. In such a case, the indirection-id references to a Segment Routing Binding SID, while the indirection-id type references the Bindging SID semantic. 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 with an explicit path and can be setup by a controller using BGP as specified in [I-D.sreekantiah-idr-segment-routing-te] [5] or using PCE as detailed in draft-ietf-pce-segment-routing [7]. When a BGP speaker receives a flow-spec route with a 'redirect to Binding SID' extended community, it installs a traffic filtering rule that matches the packets described by the NLRI field and redirects them to the explicit path associated

with the Binding SID. The explicit path is specified as a set/stack of segment identifiers as detailed in the previous documents. The stack of segment identifiers is now imposed on packets matching the flow-spec rule to perform redirection as per the explicit path setup prior. The encoding of the Binding SID value is specified in section 4, with the indirection-id field now encoding the associated value for the binding SID.

Requirements:

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

3.3. Redirection to Next-next-hop tunnels

Possible Indirection-ID type examples:

- When deploying on flowspec client using a localised Indirection-id table the TID (Table ID) is used: one indirection-id community of type 0 (localised ID) with TID=0 and second indirection-id community of type 0 with TID=1

Description:

A Third use-case is when a BGP Flowspec controller sends a single flowspec policy route to flowspec clients to signal redirection towards next-next-hop tunnels. In this use-case The flowspec rule is instructing the Flowspec client to redirect traffic using a sequence of indirection-id extended communities. The sequence of indirection-ids is managed using Tunnel IDs (TID).

Segment Routing Example:

i.e. a classic Segment Routing example would be DDoS mitigation towards a Segment Routing Central Egress Path Engineered tunnel [4]. To steer DDoS traffic towards egress peer engineering paths, a first indirection-id (i.e. TID=0) will steer traffic onto a tunnel to an egress router, while the second indirection-id (TID=1) is used steer the egress router arrived traffic onto a pre-identified interface/ peer. The flowspec client will for this use-case in the simpliest implementation dynamically append 2 MPLS labels. A first MPLS label (the outer label) is used to steer the original packet to the egress node, while the next MPLS label (the inner label, corresponding with the indirection-id identified with TID=1) instructs the egress router to steer the original packet to a pre-defined interface/peer corresponding principles documented by [4].

Requirements:

To achieve this type of redirection to next-next-hop tunnels, for each flowspec route, multiple indirection-ids, each using a unique Tunnel ID are imposed upon a the flowspec policy rule. It is required that there is synchronisation between the labels used by the Egress Peer Engineering egress router and the flowspec client originally imposing the sequens of EPE Segment Routing segments. It is required that the the engineered next-next-hop tunnel MUST be operational and allow packets to be exchanged between tunnel head- and tail-end.

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.

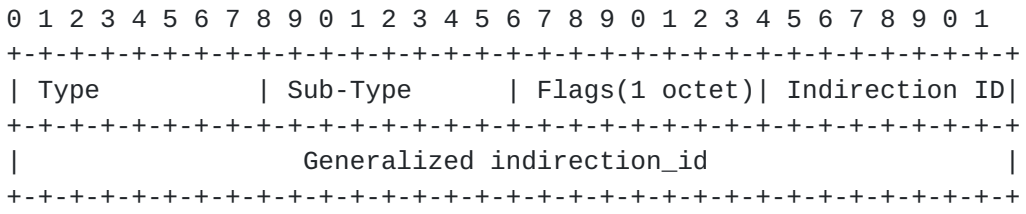

Figure 1

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.

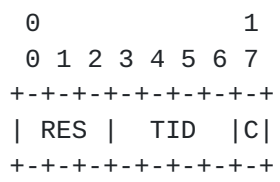

Figure 2

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.

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)

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

3 - AS (Autonomous System) ID (The flowspec client uses the received indirection-id as a Segment Routing Autonomous System ID to redirect traffic towards)

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

5 - Multicast ID (The flowspec client uses the received indirection-id as a Segment Routing Multicast 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]

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

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

9 - ECMP (Equal Cost Multi-Path) ID (The flowspec client uses the received indirection-id as a Segment Routing PeerSet ID to redirect traffic towards)

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

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

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

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

5. Redirect using localised indirection-id mapping table

When a BGP speaker receives a flowspec policy route with a 'redirect to indirection-id' extended community and this route represents the one and only best path or an equal cost multipath, 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 indirection-id local recursed path. To construct the local recursed path, the flowspec client does a local indirection-id mapping table lookup (i.e. indirection-id type = 0) using the key comprised of the indirection-id 32-bit value and indirection-id type (=0) to retrieve the correct redirection information.

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]
2	Agency ID	[RFC-To-Be]
3	AS (Autonomous System) ID	[RFC-To-Be]
4	Anycast ID	[RFC-To-Be]
5	Multicast ID	[RFC-To-Be]
6	Tunnel ID (Tunnel Binding ID)	[RFC-To-Be]
7	VPN ID	[RFC-To-Be]
8	OAM ID	[RFC-To-Be]
9	ECMP (Equal Cost Multi-Path) ID	[RFC-To-Be]
10	QoS ID	[RFC-To-Be]
11	Bandwidth-Guarantee ID	[RFC-To-Be]
12	Security ID	[RFC-To-Be]
13	Multi-Topology 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, <http://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
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: keyupate@cisco.com

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

Email: lizhenbin@huawei.com