

Routing Area Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 28, 2017

S. Litkowski
Orange
A. Simpson
Nokia
K. Patel
Arrcus, Inc
J. Haas
Juniper Networks
L. Yong
Huawei
October 25, 2016

**Applying BGP flowspec rules on a specific interface set
draft-ietf-idr-flowspec-interfaceset-02**

Abstract

BGP flowspec is an extension to BGP that allows for the dissemination of traffic flow specification rules. The primary application of this extension is DDoS mitigation where the flowspec rules are applied in most cases to all peering routers of the network.

This document will present another use case of BGP flowspec where flow specifications are used to maintain some access control lists at network boundary. BGP flowspec is a very efficient distributing machinery that can help in saving OPEX while deploying/updating ACLs. This new application requires flow specification rules to be applied only on a specific subset of interfaces and in a specific direction.

The current specification of BGP flowspec ([\[RFC5575\]](#)) introduces the notion of flow specification (which describes the matching criterion) and traffic filtering actions. The flow specification is encoded as part of the NLRI while the traffic filtering actions are encoded as extended communities. The combination of a flow specification and one or more actions is known as a flow specification rule. [\[RFC5575\]](#) does not detail where the flow specification rules need to be applied.

Besides the flow specification and traffic filtering actions, this document introduces the notion of traffic filtering scope in order to drive where a particular rule must be applied. In particular, this document introduces the "interface-set" traffic filtering scope that could be used in parallel of traffic filtering actions (marking, rate-limiting ...). The purpose of this extension is to inform remote routers about groups of interfaces where the rule must be applied.

This extension can also be used in a DDoS mitigation context where a provider wants to apply the filtering only on specific peers.

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 [[RFC2119](#)].

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 April 28, 2017.

Copyright Notice

Copyright (c) 2016 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.

Table of Contents

1.	Use case	3
1.1.	Specific filtering for DDoS	3
1.2.	ACL maintenance	4
2.	Collaborative filtering and managing filter direction	5
3.	Traffic filtering scope	6
4.	Interface specific filtering using BGP flowspec	7
5.	Interface-set extended community	8
6.	Handling rules from different sources in the processing pipe	9
6.1.	Combining Policy Based Routing with BGP flowspec and interface-set	10
6.2.	Combining flow collection with BGP flowspec and interface-set	11
7.	Scaling of per interface rules	11
8.	Deployment considerations	12
9.	Security Considerations	12
10.	Acknowledgements	12
11.	IANA Considerations	13
12.	References	13
12.1.	Normative References	13
12.2.	Informative References	14
	Authors' Addresses	14

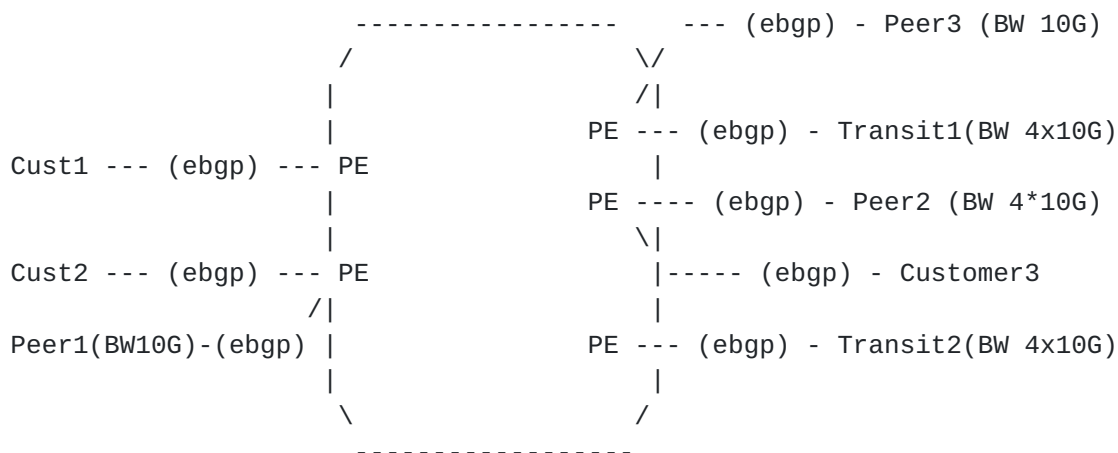
[1.](#) Use case[1.1.](#) Specific filtering for DDoS

Figure 1

The figure 1 above displays a typical service provider Internet network owing Customers, Peers and Transit. To protect pro actively against some attacks (e.g. DNS, NTP ...), the service provider may want to deploy some rate-limiting of some flows on peers and transit

links. But depending on link bandwidth, the provider may want to apply different rate-limiting values.

For 4*10G links peer/transit, it may want to apply a rate-limiting of DNS flows of 1G, while on 10G links, the rate-limiting would be set to 250Mbps. Customer interfaces must not be rate-limited.

BGP flowspec infrastructure may already be present on the network, and all PEs may have a BGP session running flowspec address family. The flowspec infrastructure may be reused by the service provider to implement such rate-limiting in a very quick manner and being able to adjust values in future quickly without having to configure each node one by one. Using the current BGP flowspec specification, it would not be possible to implement different rate limiter on different interfaces of a same router. The flowspec rule is applied to all interfaces in all directions or on some interfaces where flowspec is activated but flowspec rule set would be the same among all interfaces.

Section [Section 4](#) will detail a solution to address this use case using BGP flowspec.

1.2. ACL maintenance

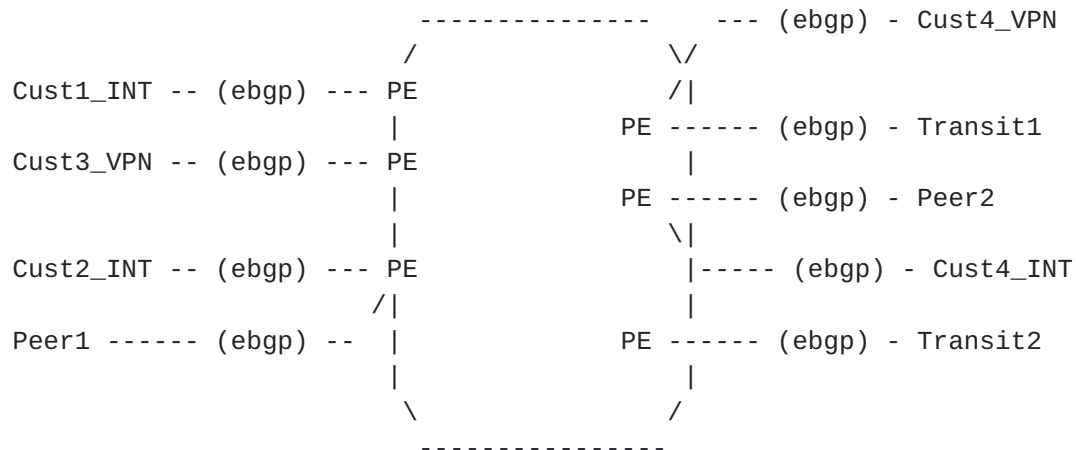


Figure 2

The figure 1 above displays a typical service provider multiservice network owing Customers, Peers and Transit for Internet, as well as VPN services. The service provider requires to ensure security of its infrastructure by applying ACLs at network boundary. Maintaining and deploying ACLs on hundreds/thousands of routers is really painful and time consuming and a service provider would be interested to deploy/updates ACLs using BGP flowspec. In this scenario, depending

on the interface type (Internet customer, VPN customer, Peer, Transit ...) the content of the ACL may be different.

We foresee two main cases :

- o Maintaining complete ACLs using flowspec : in this case all the ingress ACL are maintained and deployed using BGPflowspec. See section [Section 9](#) for more details on security aspects.
- o Requirement of a quick deployment of a new filtering term due to a security alert : new security alerts often requires a fast deployment of new ACL terms. Using traditional CLI and hop by hop provisioning, such deployment takes time and network is unprotected during this time window. Using BGP flowspec to deploy such rule, a service provider can protect its network in few seconds. Then the SP can decide to keep the rule permanently in BGP flowspec or update its ACL or remove the entry (in case equipments are not vulnerable anymore).

Section [Section 4](#) will detail a solution to address this use case using BGP flowspec.

2. Collaborative filtering and managing filter direction

[RFC5575] states in [Section 5](#). : "This mechanism is primarily designed to allow an upstream autonomous system to perform inbound filtering in their ingress routers of traffic that a given downstream AS wishes to drop."

In case of networks collaborating in filtering, there is a use case for performing outbound filtering. Outbound filtering allows to apply traffic action one step before and so may allow to prevent impact like congestions.

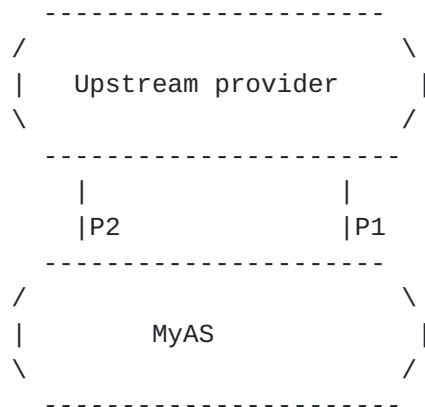


Figure 3

In the figure above, MyAS is connected to an upstream provider. If a malicious traffic comes in from the upstream provider, it may congestion P1 or P2 links. If MyAS apply inbound filtering on P1/P2 using BGP flowspec, the congestion issue will not be solved.

Using collaborative filtering, the upstream provider may propose to MyAS to filter malicious traffic sent to it. We propose to enhance [\[RFC5575\]](#) to make myAS able to send BGP flowspec updates (on eBGP sessions) to the upstream provider to request outbound filtering on peering interfaces towards MyAS. When the upstream provider will receive the BGP flowspec update from MyAS, the BGP flowspec update will contain request for outbound filtering on a specific set of interfaces. The upstream provider will apply automatically the requested filter and congestion will be prevented.

3. Traffic filtering scope

We see with the use case described above that some BGP flowspec rules may need to be applied only on specific elements of the network (interfaces, nodes ...). The basic specification detailed in [\[RFC5575\]](#) does not address this and does not give any detail on where the traffic filtering rule need to be applied.

In addition to the flow specification (flow matching criterion) and traffic filtering actions described in [\[RFC5575\]](#), this document introduces the notion of traffic filtering scope. The traffic filtering scope will describe where a particular flow specification rule must be applied.

Using a traffic filtering scope in a BGP flowspec rule is optional. When a rule does not contain any traffic filtering scope parameter, [\[RFC5575\]](#) applies.

4. Interface specific filtering using BGP flowspec

The use case detailed above requires application of different BGP flowspec rules on different set of interfaces.

We propose to introduce, within BGP flowspec, a traffic filtering scope that identifies a group of interfaces where a particular filter should apply on. Identification of interfaces within BGP flowspec will be done through group identifiers. A group identifier marks a set of interfaces sharing a common administrative property. Like a BGP community, the group identifier itself does not have any significance. It is up to the network administrator to associate a particular meaning to a group identifier value (e.g. group ID#1 associated to Internet customer interfaces). The group identifier is a local interface property. Any interface may be associated with one or more group identifiers using manual configuration.

When a filtering rule advertised through BGP flowspec must be applied only to particular sets of interfaces, the BGP flowspec BGP update will contain the identifiers associated with the relevant sets of interfaces. In addition to the group identifiers, it will also contain the direction the filtering rule must be applied in (see [Section 5](#)).

Configuration of group identifiers associated to interfaces may change over time. An implementation MUST ensure that the filtering rules (learned from BGP flowspec) applied to a particular interface are always updated when the group identifier mapping is changing.

Considering figure 2, we can imagine the following design :

- o Internet customer interfaces are associated with group-identifier 1.
- o VPN customer interfaces are associated with group-identifier 2.
- o All customer interfaces are associated with group-identifier 3.
- o Peer interfaces are associated with group-identifier 4.
- o Transit interfaces are associated with group-identifier 5.
- o All external provider interfaces are associated with group-identifier 6.
- o All interfaces are associated with group-identifier 7.

If the service provider wants to deploy a specific inbound filtering on external provider interfaces only, the provider can send the BGP flow specification using group-identifier 6 and including inbound direction.

There are some cases where nodes are dedicated to specific functions (Internet peering, Internet Edge, VPN Edge, Service Edge ...), in this kind of scenario, there is an interest for a constrained distribution of filtering rules that are using the interface specific filtering. Without the constrained route distribution, all nodes will received all the filters even if they are not interested in those filters. Constrained route distribution of flowspec filters would allow for a more optimized distribution.

5. Interface-set extended community

This document proposes a new BGP Route Target extended community called "flowspec interface-set". This document so expands the definition of the Route Target extended community to allow a new value of high order octet (Type field) to be TBD (in addition to the values specified in [RFC4360](#)).

In order to ease intra-AS and inter-AS use cases, this document proposes to have a transitive as well as a non transitive version of this extended community.

This new BGP Route Target extended community is encoded as follows :

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|  Type (TBD)   |      0x02      |  Autonomous System Number   :
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
:      AS Number (cont.)           |0|I|      Group Identifier   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

The flags are :

- o 0 : if set, the flow specification rule MUST be applied in outbound direction to the interface set referenced by the following group-identifier.
- o I : if set, the flow specification rule MUST be applied in input direction to the interface set referenced by the following group-identifier.

Both flags can be set at the same time in the interface-set extended community leading to flow rule to be applied in both directions. An interface-set extended community with both flags set to zero MUST be treated as an error and as consequence, the flowspec update MUST be discarded. As having no direction indicated as no sense, there is no need to propagate the filter informations in the network.

The Group Identifier is coded as a 14-bit number (values goes from 0 to 16383).

Multiple instances of the interface-set community may be present in a BGP update. This may appear if the flow rule need to be applied to multiple set of interfaces.

Multiple instances of the community in a BGP update MUST be interpreted as a "OR" operation : if a BGP update contains two interface-set communities with group ID 1 and group ID 2, the filter would need to be installed on interfaces belonging to Group ID 1 or Group ID 2.

As using a Route Target, route distribution of flowspec NLRI with interface-set may be subject to constrained distribution as defined in [[RFC4684](#)]. Constrained route distribution for flowspec routes using interface-set requires discussion and will be addressed in a future revision of the document.

6. Handling rules from different sources in the processing pipe

A packet on a router may be processed by multiple rules that are originated by different sources (e.g. statically configured, I2RS ephemeral, BGP ephemeral ...). [[RFC5575](#)] does not provide any guidance regarding the precedence of BGP flowspec rules compared to other sources. A new version of BGP flowspec [[I-D.hares-idr-flowspec-v2](#)] should address these precedence rules definition in future.

This document only addresses the usage of "interface-set" in the framework of [[RFC5575](#)] and the following generic rules SHOULD be used by an implementation:

- o An inbound flowspec rule using interface-set SHOULD be processed after all existing inbound traffic processing rules (ACL, PBR, QoS, flow collection ...) and SHOULD be applied before forwarding decision is made.
- o An outbound flowspec rule using interface-set SHOULD be processed after all existing outbound traffic processing rules (ACL, PBR,

QoS, flow collection ...) and SHOULD be applied after forwarding decision has been made.

```
Inbound processing      Forwarding      Outbound processing
ACL->PBR->BGP_FS in    ->  Lookup    ->  PBR->QoS->I2RS->BGP_FS out
```

Example of packet processing pipe

In the example above, any BGP flowspec rule ([\[RFC5575\]](#)) with an inbound interface-set is processed after all existing features (ACL, PBR and QoS) but before the lookup is done. The outbound BGP flowspec rules with interface-set are applied after the lookup and after all any existing feature.

This section gives two examples of combining existing features with BGP flowspec+interface-set attribute on an interface.

6.1. Combining Policy Based Routing with BGP flowspec and interface-set

In this example, a router has an already configured inbound policy on an interface IF1 with the following rules:

o Static policy IN:

- * Entry 1: from udp 10/8 to 11/8 port 53 then set dscp af11 and accept
- * Entry 2: from tcp 10/8 to 11/8 port 22 then set dscp af22 and accept
- * Entry 3: from tcp 10/8 to 11/8 port 80 then set dscp af11 and accept
- * Entry 4: from ip 10/8 to 11/8 then drop

In addition to this static inbound ACL, the router receives the following unordered BGP flowspec version 1 rules with an interface-set matching IF1.

o flowspec rules IN :

- * from udp 10.0.0.1/32 to 11/8 port 53 then drop
- * from tcp 10/8 to 11/8 port 80 then set dscp af32 and accept
- * from udp 10/8 to 11/8 then accept

The combination of the static inbound ACL and the inbound "interface-set" flowspec rules should result in the following packet processing:

- o a UDP flow from 10.0.0.1 to 11.0.0.2 on port 53 will be rejected: firstly, it is allowed by the static ACL and DSCP is set to af11 but then it is dropped by the flowspec filter.
- o a UDP flow from 10.0.0.2 to 11.0.0.2 on port 53 will be accepted and DSCP will be set to af11.
- o a TCP flow from 10.0.0.2 to 11.0.0.2 on port 80 will be accepted and DSCP will be set to af32: firstly, the static PBR rule set the DSCP to af11 and accepts the packet, then the flowspec filter rewrites the DSCP to af32 and accepts it also.

6.2. Combining flow collection with BGP flowspec and interface-set

A router may activate flow collection features (used in collaboration with Netflow export). Flow collection can be done at input side or output side. When a router is configured to collect flow informations on an inbound interface, the flow collection happens before any BGP flowspec rule with interface-set: if a particular packet flow is denied by a BGP flowspec rule, it will still be collected.

The same happens if a router is configured to collect flow informations on an outbound interface, the flow collection happens, and then the BGP flowspec rule is applied: the flow is collected whatever the result of the BGP flowspec rule.

7. Scaling of per interface rules

Without "interface-set", all the interfaces are using the same flowspec filter, while with "interface-set" different interfaces may have different flowspec filters (with different terms and actions). Having such flowspec rules that are applied on specific interfaces may create forwarding instructions that may be harder to share within the forwarding plane: a particular term may be present or not in the filter of a particular interface.

An implementation SHOULD take care about trying to keep sharing forwarding structures as much as possible in order to limit the scaling impact. How the implementation would do so is out of scope of the document.

8. Deployment considerations

There are some cases where a particular BGP flowspec NLRI may be advertised to different interface groups with a different action. For example, a service provider may want to discard all ICMP traffic from customer interfaces to infrastructure addresses and want to rate-limit the same traffic when it comes from some internal platforms. These particular cases require ADD-PATH to be deployed in order to ensure that all paths (NLRI+interface-set group-id+actions) are propagated within the BGP control plane. Without ADD-PATH, only a single "NLRI+interface-set group-id+actions" will be propagated, so some filtering rules will never be applied.

9. Security Considerations

Managing permanent Access Control List by using BGP flowspec as described in [Section 1.2](#) helps in saving roll out time of such ACL. However some ACL especially at network boundary are critical for the network security and loosing the ACL configuration may lead to network open for attackers.

By design, BGP flowspec rules are ephemeral : the flow rule exists in the router while the BGP session is UP and the BGP path for the rule is valid. We can imagine a scenario where a Service Provider is managing the network boundary ACLs by using only flowspec. In this scenario, if , for example, an attacker succeed to make the internal BGP session of a router to be down , it can open all boundary ACLs on the node, as flowspec rules will disappear due to the BGP session down.

In reality, the chance for such attack to occur is low, as boundary ACLs should protect the BGP session from being attacked.

In order to complement the BGP flowspec solution is such deployment scenario and provides security against such attack, a service provider may activate Long lived Graceful Restart [[I-D.uttaro-idr-bgp-persistence](#)] on the BGP session owning flowspec address family. So in case of BGP session to be down, the BGP paths of flowspec rules would be retained and the flowspec action will be retained.

10. Acknowledgements

Authors would like to thanks Wim Hendrickx and Robert Raszuk for their valuable comments.

11. IANA Considerations

This document requests a new type from the "BGP Transitive Extended Community Types" extended community registry. This type name shall be 'FlowSpec'.

This document requests a new type from the "BGP Non-Transitive Extended Community Types" extended community registry. This type name shall be 'FlowSpec'.

This document requests creation of a new registry called "FlowSpec Extended Community Sub-Types". This registry contains values of the second octet (the "Sub-Type" field) of an extended community when the value of the first octet (the "Type" field) is to one of those allocated in this document.

Within this new registry, this document requests a new subtype (suggested value 0x02), this sub-type shall be named "interface-set".

12. References

12.1. Normative References

[I-D.ietf-idr-rtc-no-rt]

Rosen, E., Patel, K., Haas, J., and R. Raszuk, "Route Target Constrained Distribution of Routes with no Route Targets", [draft-ietf-idr-rtc-no-rt-05](#) (work in progress), May 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://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, <<http://www.rfc-editor.org/info/rfc4360>>.

[RFC4684] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", [RFC 4684](#), DOI 10.17487/RFC4684, November 2006, <<http://www.rfc-editor.org/info/rfc4684>>.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", [RFC 5575](https://www.rfc-editor.org/info/rfc5575), DOI 10.17487/RFC5575, August 2009, <<http://www.rfc-editor.org/info/rfc5575>>.

12.2. Informative References

[I-D.hares-idr-flowspec-v2]
Hares, S., "BGP Flow Specification Version 2", [draft-hares-idr-flowspec-v2-00](#) (work in progress), June 2016.

[I-D.uttaro-idr-bgp-persistence]
Uttaro, J., Chen, E., Decraene, B., and J. Scudder, "Support for Long-lived BGP Graceful Restart", [draft-uttaro-idr-bgp-persistence-03](#) (work in progress), November 2013.

Authors' Addresses

Stephane Litkowski
Orange

Email: stephane.litkowski@orange.com

Adam Simpson
Nokia

Email: adam.simpson@nokia.com

Keyur Patel
Arrcus, Inc

Email: keyur@arrcus.com

Jeff Haas
Juniper Networks

Email: jhaas@juniper.net

Lucy Yong
Huawei

Email: lucy.yong@huawei.com

