Routing Area Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 10, 2016

S. Litkowski
Orange
A. Simpson
Alcatel Lucent
K. Patel
Cisco
J. Haas
Juniper Networks

December 8, 2015

Applying BGP flowspec rules on a specific interface set
draft-litkowski-idr-flowspec-interfaceset-03

Abstract

BGP Flow-spec 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 Flow-spec 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 Flow-spec does not detail where the flow specification rules need to be applied.

This document presents a new interface-set flowspec action that will be used in complement of other actions (marking, rate-limiting ...). The purpose of this extension is to inform remote routers on where to apply the flow specification.

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

Copyright Notice

Table of Contents

1. Use case

1.1. Specific filtering for DDoS

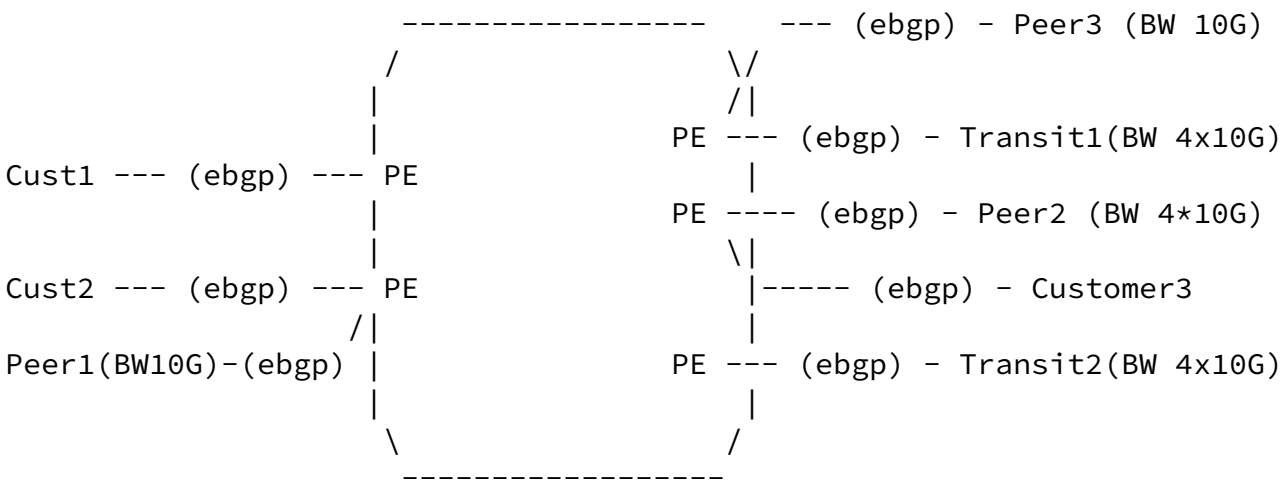


Figure 1

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 Flow-spec 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 3 will detail a solution to address this use case using BGP Flowspec.

1.2. ACL maintenance

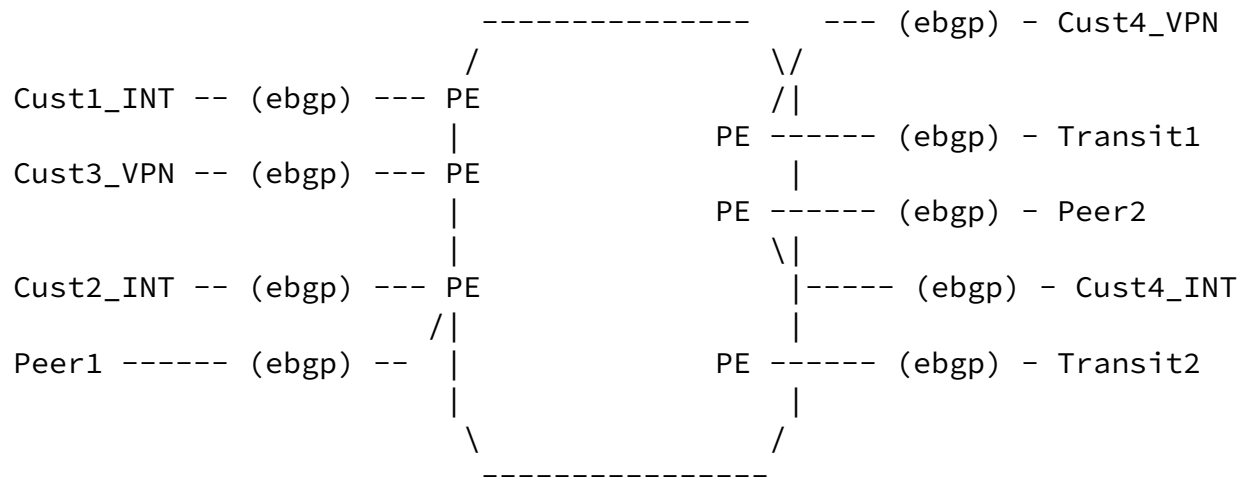


Figure 2

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 :

- Maintaining complete ACLs using flowspec : in this case all the ingress ACL are maintained and deployed using BGPFlowspec. See section Section 8 for more details on security aspects.

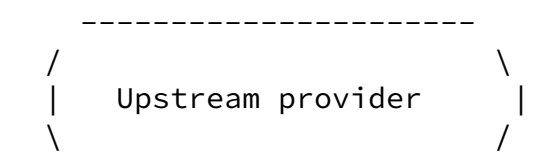
- Requirement of a quick deployment of a new filtering term due to a security alert : new security alerts often requires a fast deployment of new ACL terms. Using traditional CLI and hop by hop provisioning, such deployment takes time and network is unprotected during this time window. Using BGP flowspec 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 3 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.



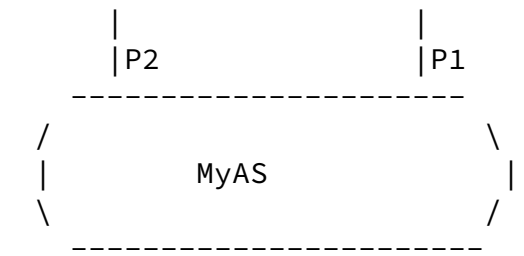


Figure 3

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. Interface specific filtering using BGP flowspec

The use case detailed above requires application of different BGP Flowspec rules on different set of interfaces. The basic specification detailed in [RFC5575] does not address this and does not give any detail on where the FlowSpec filter need to be applied.

We propose to introduce, within BGP Flowspec, an identification 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 4).

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 :

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

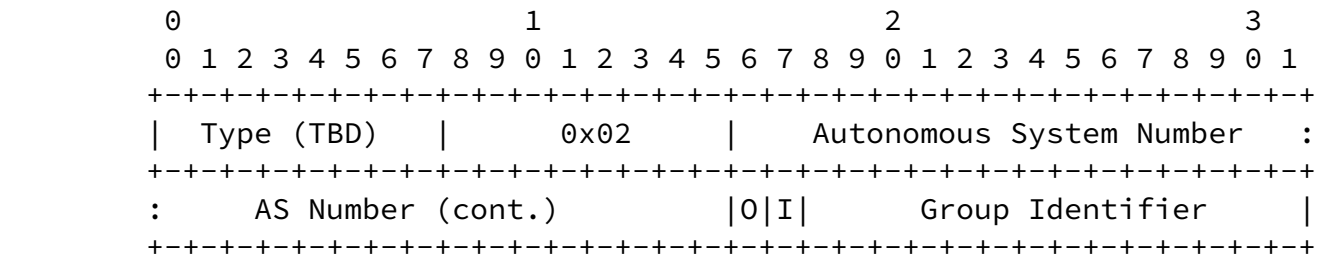
distribution of filtering rules that are using the interface specific filtering. Without the constrained route distribution, all nodes will received all the filters even if they are not interested in those filters. Constrained route distribution of flowspec filters would allow for a more optimized distribution.

4. 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 :



The flags are :

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

- 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.

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.

5. Interaction with permanent traffic actions

[RFC5575] states that BGP Flowspec is primarily designed to allow upstream AS to perform inbound filtering in their ingress routers. This specification does not precise where this ingress filtering should happen in the packet processing pipe.

This proposal enhances [RFC5575] in order to add action on traffic coming from or going to specific interfaces. Based on this enhancement, some new requirements come to implementations.

An implementation SHOULD apply input actions (I bit set) within the input packet processing pipe. An implementation SHOULD apply output actions (O bit set) within the output packet processing pipe.

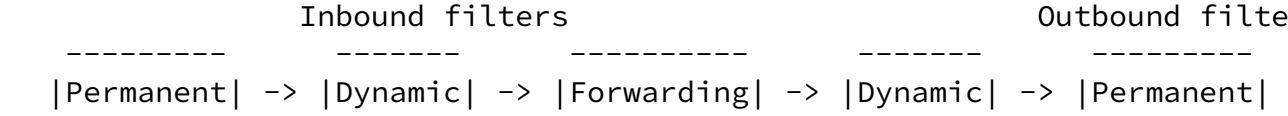
As input and output processing pipes may also involve already present static/permanent features that will manipulate the packet, the next sections will try to clarify how the static behaviors should interact will BGP flowspec actions.

5.1. Interaction with interface ACLs

Deploying interface specific filters using BGP FlowSpec (dynamic entries) may interfere with existing permanent interface ACL (static entries). The content of the existing permanent ACL MUST NOT be altered by dynamic entries coming from BGP FlowSpec. Permanent ACLs are using a specific ordering which is not compatible with the ordering of FS rules and misordering of ACL may lead to undesirable behaviour. In order, to keep a deterministic and well known behaviour, an implementation SHOULD process the BGP FlowSpec ACL as follows :

- In inbound direction, the permanent ACL action is applied first followed by FlowSpec action. This gives the primary action to the permanent ACL as it is done today.

- In outbound direction, FlowSpec action action is applied first followed by permanent ACL. This gives the final action to the permanent ACL as it is done today.



In order for a flow to be accepted, the flow must be accepted by the two ACLs and a flow is rejected when one of the ACL rejects it as described in the table below :

Permanent ACL entry action	FlowSpec ACL entry action	Result action
Drop	Drop	Drop
Drop	Accept	Drop
Accept	Drop	Drop
Accept	Accept	Accept

Example :

- ACL permanent IN :

  * Entry 1 : permit udp from 10/8 to 11/8 port 53

* Entry 2 : permit tcp from 10/8 to 11/8 port 22

* Entry 3 : deny ip from 10/8 to 11/8

- ACL dynamic FlowSpec IN :

  * Entry 1 : deny udp from 10.0.0.1/32 to 11/8 port 53

  * Entry 2 : permit tcp from 10/8 to 11/8 port 80

In the example above :

- a UDP flow from 10.0.0.1 to 11.0.0.2 on port 53 will be rejected because the dynamic ACL rejects it.

- a UDP flow from 10.0.0.2 to 11.0.0.2 on port 53 will be accepted because both ACLs accept it.

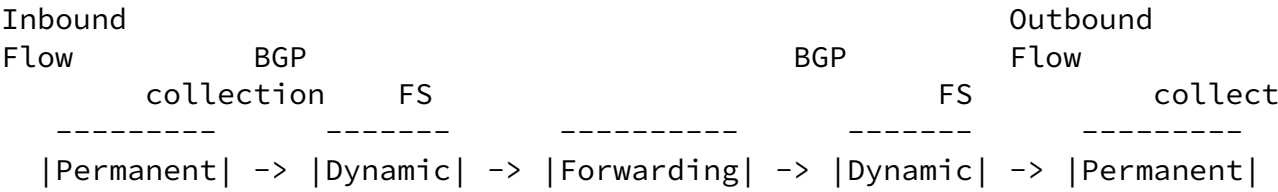
- a TCP flow from 10.0.0.2 to 11.0.0.2 on port 80 will be rejected because permanent ACL rejects it.

5.2. Interaction with flow collection

A router may activate flow collection features (used in collaboration with Netflow export). Flow collection can be done at input side or output side. As for ACL, an implementation SHOULD process :

- BGP FS rules after the inbound flow collection : in case of DDoS protection, it is important to keep monitoring of attack flows and so performing action, after collection.

- BGP FS rules before the outbound flow collection : purpose of outbound flow collection is really to track flows that are exiting the interface. BGP FS rules should not interfere in this.



6. Scaling of per interface rules

Creating rules that are applied on specific interfaces may create forwarding rules that may be harder to share.

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.

7. 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 group+actions) are propagated within the BGP control plane. Without ADD-PATH, only a single "NLRI+interface group+actions" will be propagated, so some filtering rules will never be applied.

8. 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.

9. Acknowledgements

Authors would like to thanks Wim Hendrickx for his valuable comments.

10. 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".

11. References

11.1. 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-04 (work in progress), November 2015.

[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, DOI 10.17487/RFC5575, August 2009, <http://www.rfc-editor.org/info/rfc5575>.

11.2. Informative References

[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
Alcatel Lucent

Email: adam.simpson@alcatel-lucent.com

Keyur Patel
Cisco

Email: keyupate@cisco.com

Jeff Haas
Juniper Networks

Email: jhaas@juniper.net