Idr Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 1, 2016

Q. Liang
J. You
Huawei
R. Raszuk
Nozomi
D. Ma
Cisco Systems

September 29, 2015

Carrying Label Information for BGP FlowSpec
draft-liang-idr-bgp-flowspec-label-01

Abstract

This document specifies a method in which the label mapping information for a particular FlowSpec rule is piggybacked in the same Border Gateway Protocol (BGP) Update message that is used to distribute the FlowSpec rule. Based on the proposed method, the Label Switching Routers (LSRs) (except the ingress LSR) on the Label Switched Path (LSP) can use label to indentify the traffic matching a particular FlowSpec rule; this facilitates monitoring and traffic statistics for FlowSpec rules.

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 1, 2016.

Table of Contents

1. Introduction 2
2. Terminology 4
3. Protocol Extensions 4
4. IANA Considerations 6
5. Security considerations 6
6. Acknowledgement 6
7. Normative References 6
Authors' Addresses 7

1. Introduction

[RFC5575] defines the flow specification (FlowSpec) that is an n-tuple consisting of several matching criteria that can be applied to IP traffic. The matching criteria can include elements such as source and destination address prefixes, IP protocol, and transport protocol port numbers. A given IP packet is said to match the defined flow if it matches all the specified criteria. [RFC5575] also defines a set of filtering actions, such as rate limit, redirect, marking, associated with each flow specification. A new Border Gateway Protocol Network Layer Reachability Information (BGP NLRI) (AFI/SAFI: 1/133 for IPv4, AFI/SAFI: 1/134 for VPNv4) encoding format is used to distribute traffic flow specifications.

[RFC3107] specifies the way in which the label mapping information for a particular route is piggybacked in the same Border Gateway Protocol Update message that is used to distribute the route itself. Label mapping information is carried as part of the Network Layer Reachability Information (NLRI) in the Multiprotocol Extensions attributes. The Network Layer Reachability Information is encoded as one or more triples of the form <length, label, prefix>. The NLRI

contains a label is indicated by using Subsequent Address Family Identifier (SAFI) value 4.

[RFC4364] describes a method in which each route within a Virtual Private Network (VPN) is assigned a Multiprotocol Label Switching (MPLS) label. If the Address Family Identifier (AFI) field is set to 1, and the SAFI field is set to 128, the NLRI is an MPLS-labeled VPN-IPv4 address.

In BGP VPN/MPLS networks, when FlowSpec rules on multiple forwarding devices in the network bound with labels form one or more LSPs, only the ingress LSR (Label Switching Router) needs to identify a particular traffic flow based on the matching criteria and then steers the packet to a corresponding LSP (Label Switched Path). Other LSRs of the LSP just need to forward the packet according to the label carried in it.

Though the FlowSpec rule could use the label(s) bound with the best-match unicast route for the destination prefix embedded in the FlowSpec rule or the best-match route to the target IP in the 'redirect to IP' action, this way means that if two or more FlowSpec rules have the same best-match unicast route for the embedded destination prefix or the same best-match route to target IP in the 'redirect to IP' action; they would be mapped to the same label. This would affect monitoring and traffic statistics facilities, because each FlowSpec rule requires an independent statistic and log data, which is described in Section 9 [RFC5575]. The LSRs (except the ingress LSR) on the LSP can use label to indentify the traffic matching a particular FlowSpec rule; this facilitates monitoring and traffic statistics for FlowSpec rules.

So this document proposes that the BGP router supports to allocate a label to one or more FlowSpec rule(s), the forwarding path is still decided by the best-match unicast route for the embedded destination prefix or the best-match route to target IP in the 'redirect to IP' action. Figure 1 gives an example that FlowSpec rule bound with a label is disseminated in the network.

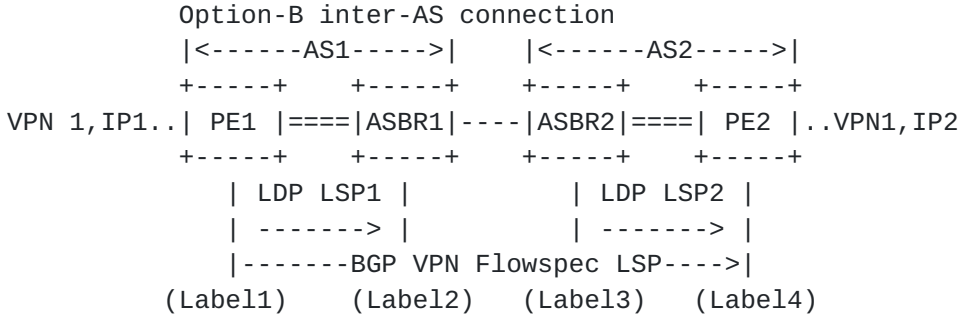


Figure 1: Usage of FlowSpec with Label

```
FlowSpec rule1 (injected in PE2):
     Filters:
        destination ip prefix:IP2/32
        source ip prefix:IP1/32
     Actions:
        traffic-marking: 1

Labels allocated for FlowSpec1:
     Label4 allocated by PE2
     Label3 allocated by ASBR2
     Label2 allocated by ASBR1
     Label1 allocated by PE1

PE2 disseminates the FlowSpec1 bound with Label4 to ASBR2.
ASBR2 disseminates the FlowSpec1 bound with Label3 to ASBR1.
ASBR1 disseminates the FlowSpec1 bound with Label2 to PE1.

Forwarding information for the traffic from IP1 to IP2 in the Routers:
     PE1: in(<IP2,IP1>) --> out(Label2)
     ASBR1: in(Label2) --> out(Label3)
     ASBR2: in(Label3) --> out(Label4)
     PE2: in(Label4) --> out(--)
```

So ASBR1 can do traffic statistics for FlowSpec rule 1 based on Label2; ASBR2 can do it based on Label3; and PE2 can do it based on Label4.

2. Terminology

This section contains definitions of terms used in this document.

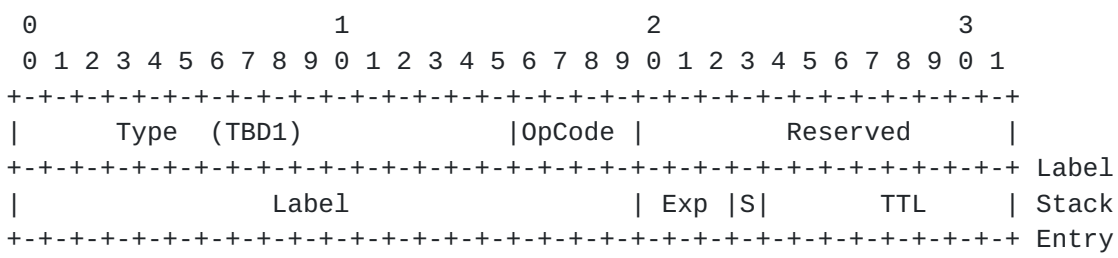
Flow Specification (FlowSpec): A flow specification is an n-tuple consisting of several matching criteria that can be applied to IP traffic, including filters and actions. Each FlowSpec consists of a set of filters and a set of actions.

3. Protocol Extensions

In this document, BGP is used to distribute the FlowSpec rule bound with label(s). A new label-action is defined as BGP extended community value based on Section 7 of [RFC5575].

type	extended community	encoding
TBD1	label-action	MPLS tag

Label-action is described below:



The use and the meaning of these fields are as follows:

Type: the same as defined in [RFC4360]

OpCode: Operation code

OpCode	Function
0	Push the MPLS tag
1	Pop the outermost MPLS tag in the packet
2	Swap the MPLS tag with the outermost MPLS tag in the packet
3~15	Reserved

When the OpCode field is set to 1, the label stack entry is invalid, and the router SHOULD pop the existing outermost MPLS tag in the packet.

When the OpCode field is set to 2, the router SHOULD swap the label stack entry with the existing outermost MPLS tag in the packet. If the packet has no MPLS tag, it just pushes the label stack entry.

The OpCode 0 or 1 may be used in some SDN networks, such as the scenario described in [I-D.filsfils-spring-segment-routing-central-epe].

The OpCode 2 can be used in traditional BGP MPLS/VPN networks.

Bottom of Stack (S): the same as defined in [RFC3032]. It SHOULD be invalid, and set to zero by default. It MAY be modified by the forwarding router locally.

Time to Live (TTL): the same as defined in[RFC3032]. It MAY be modified by the forwarding router locally.

Experimental Use (Exp): the same as defined in [RFC3032]. It MAY be modified by the forwarding router according to the local routing policy.

Label: the same as defined in [RFC3032].

A FlowSpec rule MAY include one or more ordering label-action(s). The arrival order of the label-actions decides the action order.

If the BGP router allocates a label for a FlowSpec rule and disseminates the labeled FlowSpec rule to the upstream peers, it can use the label to match the traffic identified by the FlowSpec rule in the forwarding plane.

4. IANA Considerations

For the purpose of this work, IANA should allocate value for the type of label-action:

TBD1 for label-action

5. Security considerations

This extension to BGP does not change the underlying security issues inherent in the existing BGP.

6. Acknowledgement

The authors would like to thank Shunwan Zhuang, Zhenbin Li, Peng Zhou and Jeff Haas for their comments.

7. Normative References

[I-D.filsfils-spring-segment-routing-central-epe]
Filsfils, C., Previdi, S., Patel, K., Shaw, S., Ginsburg, D., and D. Afanasiev, "Segment Routing Centralized Egress Peer Engineering", draft-filsfils-spring-segment-routing-central-epe-05 (work in progress), August 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>.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <http://www.rfc-editor.org/info/rfc3032>.

[RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", RFC 3107, DOI 10.17487/RFC3107, May 2001, <http://www.rfc-editor.org/info/rfc3107>.

[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>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <http://www.rfc-editor.org/info/rfc4364>.

[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>.

Authors' Addresses

Qiandeng Liang
Huawei
101 Software Avenue, Yuhuatai District
Nanjing, 210012
China

Email: liangqiandeng@huawei.com

Jianjie You
Huawei
101 Software Avenue, Yuhuatai District
Nanjing, 210012
China

Email: youjianjie@huawei.com

Robert Raszuk
Nozomi

Email: robert@raszuk.net

Dan Ma
Cisco Systems

Email: danma@cisco.com