Idr Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 22, 2016

Q. Liang
S. Hares
J. You
Huawei
R. Raszuk
Nozomi
D. Ma
Cisco Systems
March 21, 2016

Carrying Label Information for BGP FlowSpec
draft-liang-idr-bgp-flowspec-label-02

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 September 22, 2016.

Table of Contents

1. Introduction

This section provides the background for proposing a new action for BGP Flow specification that push/pops MPLS or swaps MPLS tags. For those familiar with BGP Flow specification ([RFC5575], [RFC7674], [I-D.ietf-idr-flow-spec-v6], [I-D.ietf-idr-flowspec-l2vpn], [I-D.ietf-idr-bgp-flowspec-oid] and MPLS ([RFC3107]) can skip this background section.

1.1. Background

[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.

1.2. MPLS Flow Specification Deployment

In BGP VPN/MPLS networks when flow specification policy rules exist on multiple forwarding devices in the network bound with labels from one or more LSPs, only the ingress LSR (Label Switching Router) needs to identify a particular traffic flow based on the matching criteria for flow. Once the flow is match by the ingress LSR, the ingress LSR 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.

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. Overview of Proposal

This document proposes adding a BGP-FS action in an extended community alters the label switch path associated with a matched flow. If the match does not have a label switch path, this action is skipped.

The BGP flow specification (BGP-FS) policy rule could match on the destination prefix and then utilize a BGP-FS action to adjust the label path associated with it (push/pop/swap tags.) Or a BGP-FS policy rule could match on any set of BGP-FS match conditions associated with a BGP-FS action that adjust the label switch path (push/pop/swap).

[I-D.yong-idr-flowspec-mpls-match] provides a match BGP-FS that may be used with this action to match and direct MPLS packets.

Example of Use:

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(--)
```

Labels allocated by flow policy process:

```
Label4 allocated by PE2
Label3 allocated by ASBR2
Label2 allocated by ASBR1
```

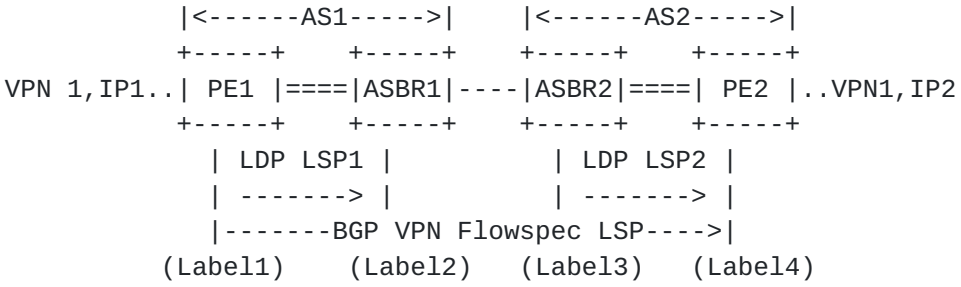


Figure 1: Usage of FlowSpec with Label

BGP-FS rule1 (locally configured):

```
Filters:
   destination ip prefix:IP2/32
   source ip prefix:IP1/32

Actions: Extended Communities
   traffic-marking: 1
   MPLS POP
```

Note:

The following Extended Communities are added/deleted

```
[rule-1a] BGP-FS action MPLS POP [used on PE2]
[rule-1b] BGP-FS action SWAP 4   [used on ASBR-2]
[rule-1c] BGP-FS action SWAP 3   [used on ASBR-1]
[rule-1d] BGP-FS action push 2   [used on PE1]
```

PE-2 Changes BGP-FS rule-1a to rule-1b prior to sending
Clears Extended Community: BGP-FS action MPLS POP
Adds Extended Community: BGP-FS action MPLS SWAP 4

ASBR-2 receives BGP-FS rule-1b (NRLI + 2 Extended Community)
Installs the BGP-FS rule-1b (MPLS SWAP 4, traffic-marking)
Changes BGP-FS rule-1b to rule-1c prior to sending to ASBR1
Clear Extended Community: BGP-FS action MPLS SWAP 4
Adds Extended Community: BGP-FS action MPLS SWAP 3

ASBR-1 Receives BGP-FS rule-1c (NLRI + 2 Extended Community)
Installs the BGP-FS rule-1c (MPLS SWAP 3, traffic-marking
Changes BGP-FS rule-1c to rule-1d prior to sending to PE-2
Clear Extended Community: BGP-FS action MPLS SWAP 3
Adds Extended Community: BGP-FS action MPLS SWAP 2

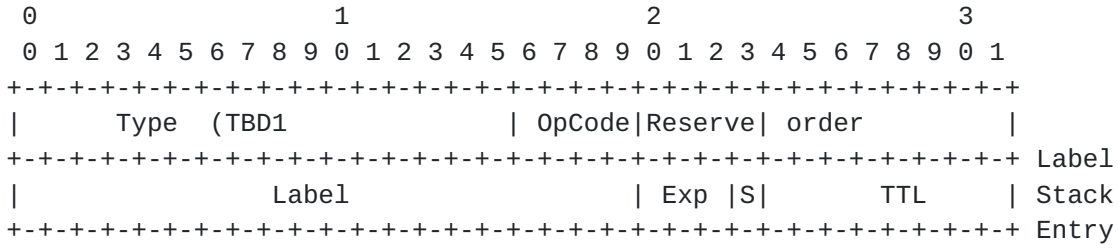
PE-1 Receives BGP-FS rule-1d (NLRI + 2 Extended Communities)
Installs BGP-FS rule-1d action [MPLS SWAP 2, traffic-marking]

4. 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 0, the label stack entry Should be pushed on the MPLS label stack.

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.

Reserved: all zeros.

Order: A FlowSpec rule MAY include one or more ordering label-action(s). If multiple label action extended communities are associated with a BGP-FS Rule, this gives the order of this in the list. The Last action received for an order will be used.

Label: the same as defined in [RFC3032].

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.

5. IANA Considerations

For the purpose of this work, IANA should allocate the following Extended community:

TBD1 for label-action

6. Security considerations

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

7. Acknowledgement

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

8. References

8.1. Normative References

[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>.

[RFC7674] Haas, J., Ed., "Clarification of the Flowspec Redirect Extended Community", RFC 7674, DOI 10.17487/RFC7674, October 2015, <http://www.rfc-editor.org/info/rfc7674>.

8.2. Informative 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.

[I-D.ietf-idr-bgp-flowspec-oid]
Uttaro, J., Filsfils, C., Smith, D., Alcaide, J., and P. Mohapatra, "Revised Validation Procedure for BGP Flow Specifications", draft-ietf-idr-bgp-flowspec-oid-02 (work in progress), January 2014.

[I-D.ietf-idr-flow-spec-v6]
McPherson, D., Raszuk, R., Pithawala, B., Andy, A., and S. Hares, "Dissemination of Flow Specification Rules for IPv6", draft-ietf-idr-flow-spec-v6-07 (work in progress), March 2016.

[I-D.ietf-idr-flowspec-l2vpn]
Weiguo, H., Litkowski, S., and S. Zhuang, "Dissemination of Flow Specification Rules for L2 VPN", draft-ietf-idr-flowspec-l2vpn-03 (work in progress), November 2015.

[I-D.yong-idr-flowspec-mpls-match]
Yong , L., Liang , Q., and J. You , "Dissemination of Flow Specification Rules for MPLS Label", March 2016.

Authors' Addresses

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

Email: liangqiandeng@huawei.com

Susan Hares
Huawei
7453 Hickory Hill
Saline, MI 48176
USA

Email: shares@ndzh.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