

Idr Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 23, 2015

Q. Liang
J. You
Huawei
October 20, 2014

BGP FlowSpec based Multi-dimensional Route Distribution
draft-liang-idr-bgp-flowspec-route-00

Abstract

This document proposes a BGP FlowSpec (Border Gateway Protocol Flow Specification) based multi-dimensional routes distribution method.

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 23, 2015.

Copyright Notice

Copyright (c) 2014 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

Internet-Draft

BGP FlowSpec

October 2014

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.	Introduction	2
1.1.	Network Load Balancing	3
1.2.	Network Diagnostics	3
1.3.	Flow-based Policy Deployment	3
1.4.	Summary	3
2.	Terminology	4
3.	BGP FlowSpec based Multi-dimensional Route	4
3.1.	Nexthop Information	4
3.2.	Carrying Label Mapping Information	5
4.	Open Issues	6
4.1.	FlowSpec Route Preference	6
4.2.	FlowSpec Route Conflict	7
4.3.	Multicast for Multi-dimensional Route	7
5.	IANA Considerations	7
6.	Security considerations	7
7.	Acknowledgement	7
8.	References	7
8.1.	Normative References	7
8.2.	Informative References	8
	Authors' Addresses	8

[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 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.

Actually, flow specification can precisely identify a particular traffic flow. It is quite appropriate to describe a multi-

dimensional route. In the following subsections, we discuss the use cases for FlowSpec based multi-dimensional route.

[1.1.](#) Network Load Balancing

Traditional routers forward traffic according to destination IP address or MAC address. The operational granularity on the traffic is coarse and limited. However, by using multi-dimensional route, multiple fields of the packet header (such as source and destination address prefixes, IP protocol, and transport protocol port numbers) can be used to identify an operational flow. For example, routers can hash the multi-dimensional routes with the same destination through different paths. Therefore, the traffic can be differentiated or managed in a finer granularity. For example, when the congestion is detected at some link, it can transfer some traffic to other links by adjusting the multi-dimensional routes. This can optimize the network load balancing.

[1.2.](#) Network Diagnostics

Traditional IP prefix or MAC based routes may aggregate the traffic that has different sources or service classes to the same link. This is difficult for traffic statistics for a particular flow in the network. However, multi-dimensional route can solve these issues such as packet loss location, traffic tracking, as it can precisely identify a particular flow.

[1.3.](#) Flow-based Policy Deployment

When using FlowSpec based multi-dimensional routes, it is more flexible and precise to deploy the flow-based policies in the network. For example, if router X wants to specify an explicit path for a particular traffic (e.g. source address = A, destination address = B), and to reserve the bandwidth through this path, this can be implemented by using multi-dimensional routes.

[1.4.](#) Summary

Especially in the data center network, campus network and service

provider networks, it's more and more important for the refined management of the network traffic.

The BGP FlowSpec based multi-dimensional route method proposed in this document can make use of the current widely deployed BGP protocol. As the multi-dimensional route usually cannot aggregate, meanwhile it has long matching key and consumes more space in the forwarding table, it is not meant to substitute the IP prefix/MAC based route. It can be regarded as a complementary routing tool for the traffic flow that having special requirements. Besides, using a dynamic multi-dimensional route distribution method can replace the traditional management approach (i.e. configuring policy-based routes

on the associated nodes in the network). It helps the automatic deployments for the policy-based routes and facilitates the route management.

Similar to IP prefix based L3 routing table, or MAC address based L2 forwarding table, FlowSpec can be regarded as a special network reachability information identity, which can be distributed by BGP. Such multi-dimensional route in this document also can be called as FlowSpec route.

[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.](#) BGP FlowSpec based Multi-dimensional Route

[3.1.](#) Nexthop Information

[RFC5575] defines a "Flow Specification" NLRI type that is encoded using MP_REACH_NLRI and MP_UNREACH_NLRI attributes as defined in [\[RFC4760\]](#). Whenever the corresponding application does not require Next-Hop information, this shall be encoded as a 0-octet length Next Hop in the MP_REACH_NLRI attribute and ignored on receipt.

However, for FlowSpec route defined in this document, nexthop information is required to be included. When included, the filtering actions [[RFC5575](#)] standardized as BGP extended community values [[RFC4360](#)] are also valid. Besides, if the FlowSpec route is the preferred one, it also determines the forwarding path of the packet that matches this FlowSpec route. Therefore, several forwarding paths from one node or multi ingress nodes to one egress node would be generated by distributing the FlowSpec routes on the forwarding devices in the network.

The nexthop information contains the network address (expressed as an IP address) of the next router on the path to the destination system. The FlowSpec routes received from a BGP peer and that are accepted in the respective Adj-RIB-In are used as input to the route selection process. The selected FlowSpec routes will be added into the FlowSpec routing table as new entries that are used to steer the packets in the data plane. Meanwhile, the routers also can do network traffic statistics based on FlowSpec routes.

[RFC5575] defines a redirect action that allows the traffic to be redirected to a VRF (virtual routing and forwarding) routing instance that lists the specified route-target in its import policy. [[I-D.ietf-idr-flowspec-redirect-ip](#)] defines a redirect-to-IP FlowSpec action that provides a method of policy-based forwarding. If a FlowSpec route carries a combination of 'redirect to IP' and 'redirect to VRF' extended communities, the BGP speaker SHOULD apply the 'redirect to IP' actions in the context of the 'target VRF', i.e. the BGP speaker redirects the matching packets or copies them towards the IPv4 or IPv6 address in the extended community's global administrator field (the 'target address').

The difference of the redirect and nexthop is that the former specifies the termination point however the latter specifies the next processing node. The redirect information is transitive but not changeable normally, i.e. the packets that match the FlowSpec routes on the different routers will be redirected to the same remote device. The nexthop information [[RFC4760](#)] is usually the local IP address of the router that disseminates this FlowSpec route, or the IP address of some remote forwarding device. The receiving router will do route iteration based on the nexthop information for obtaining the actual nexthop information. The next-hop can be

changed according to the local explicit or default route policy.

[3.2.](#) Carrying Label Mapping Information

[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 fact that the NLRI contains a label is indicated by using SAFI value 4.

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

In this document, BGP is used to distribute a particular FlowSpec route, it can also be used to distribute one or more MPLS labels that are mapped to that FlowSpec route. The Network Layer Reachability Information for FlowSpec route is encoded as one or more triples of the form <length, label, FlowSpec>, whose fields are described below.

The values for AFI/SAFI used to indicate the NLRI containing a label associated with a FlowSpec route are TBD. For example, SAFI Code (TBD1) for FlowSpec mapping label(s), and SAFI Code (TBD2) for VPN FlowSpec mapping label(s).

```
+-----+
| Length (1 octet) |
+-----+
| Label (3 octets) |
+-----+
.....
+-----+
| FlowSpec (variable) |
+-----+
```

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

Length: The Length field indicates the length in bits of the flow filters plus the label(s).

Label: The Label field carries one or more labels (that corresponds to the stack of labels [[RFC3032](#)]). Each label is encoded as 3 octets, where the high-order 20 bits contain the label value, and the low order bit contains "Bottom of Stack" [[RFC3032](#)].

FlowSpec: The FlowSpec field is the same as "flow-spec NLRI value" defined in [[RFC5575](#)]. When SAFI value is TBD1, the field contains flow filters, and when SAFI value is TBD2, the field contains a route distinguisher and flow filters.

For FlowSpec route which is associated with multiple fields of the packet header, label-based forwarding can effectively improve the route lookup performance in the data plane. When FlowSpec routes on multiple forwarding devices in the network binding the labels which makes up 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 steer 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.

[4.](#) Open Issues

[4.1.](#) FlowSpec Route Preference

The route preference for FlowSpec route is separate from the traditional IP prefix/MAC route. When the IP packet matching both

the FlowSpec route and traditional IP prefix/MAC route, then the FlowSpec route would have higher preference.

[4.2.](#) FlowSpec Route Conflict

The FlowSpec route is compliant with [[RFC5575](#)], including the constraint description for FlowSpec features. If the feature value space of the different FlowSpec routes overlaps, then it is an error.

[4.3.](#) Multicast for Multi-dimensional Route

TBD

[5.](#) IANA Considerations

The values for AFI/SAFI used to indicate the NLRI containing a label associated with a FlowSpec route need to be allocated by IANA. For example, SAFI Code (TBD1) for FlowSpec mapping label(s), and SAFI Code (TBD2) for VPN FlowSpec mapping label(s).

[6.](#) Security considerations

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

[7.](#) Acknowledgement

TBD.

[8.](#) References

[8.1.](#) Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", [RFC 3032](#), January 2001.
- [RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", [RFC 3107](#), May 2001.
- [RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", [RFC 4360](#), February 2006.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", [RFC 4364](#), February 2006.

- [RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter,

"Multiprotocol Extensions for BGP-4", [RFC 4760](#), January 2007.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", [RFC 5575](#), August 2009.

8.2. Informative References

[I-D.ietf-idr-flowspec-redirect-ip]
Uttaro, J., Haas, J., Texier, M., Andy, A., Simpson, A., and W. Henderickx, "BGP Flow-Spec Redirect to IP Action", [draft-ietf-idr-flowspec-redirect-ip-01](#) (work in progress), April 2014.

Authors' Addresses

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

Email: liuweihang@huawei.com

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

Email: youjianjie@huawei.com