IDR Working Group
W. Hao
S. Zhuang
Z. Li
Internet Draft
Huawei
Intended status: Standards Track
R.Gu
China Mobile
Expires: June 2016
December 18, 2015

Dissemination of Flow Specification Rules for NVO3
draft-hao-idr-flowspec-nvo3-03.txt

Abstract

This draft proposes a new subset of component types to support the NVO3 flow-spec application.

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/1id-abstracts.html

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

Copyright Notice

Table of Contents

1. Introduction

BGP Flow-spec is an extension to BGP that allows for the dissemination of traffic flow specification rules. It leverages the BGP Control Plane to simplify the distribution of ACLs, new filter rules can be injected to all BGP peers simultaneously without changing router configuration. The typical application of BGP Flow-spec is to automate the distribution of traffic filter lists to routers for DDOS mitigation.

RFC5575 defines a new BGP Network Layer Reachability Information (NLRI) format used to distribute traffic flow specification rules. NLRI (AFI=1, SAFI=133)is for IPv4 unicast filtering. NLRI (AFI=1, SAFI=134)is for BGP/MPLS VPN filtering. [IPv6-FlowSpec] and [Layer2-FlowSpec] extend the flow-spec rules for IPv6 and layer 2 Ethernet packets respectively. All these flow specifications match parts only reflect single layer IP/Ethernet information like source/destination MAC, source/destination IP prefix, protocol type, ports, and etc.

In cloud computing era, multi-tenancy has become a core requirement for data centers. Since NVO3 can satisfy multi-tenancy key requirements, this technology is being deployed in an increasing

number of cloud data center network. NVO3 is an overlay technology, VXLAN and NVGRE are two typical NVO3 encapsulations. GENEVE [draft-ietf-nvo3-geneve-00],GUE[draft-ietf-nvo3-gue-01] and GPE [draft-ietf-nvo3-vxlan-gpe-00] are three emerging NVO3 encapsulations in progress.

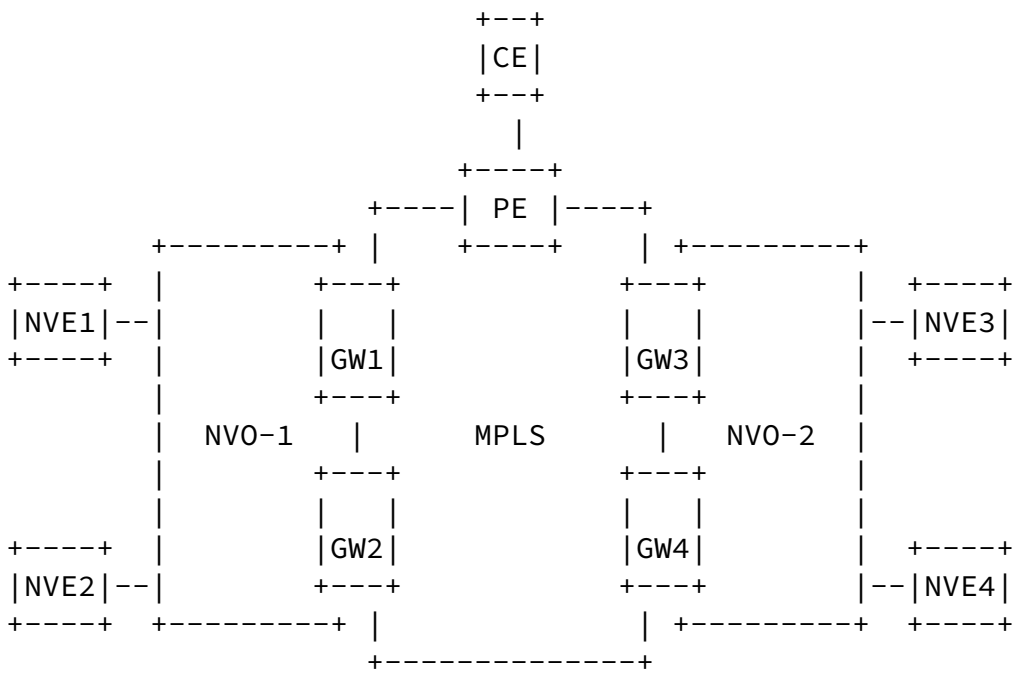


Figure 1 NVO3 data center interconnection

The MPLS L2/L3 VPN in the WAN network can be used for NVO3 based data center network interconnection. When the DC and the WAN are operated by the same administrative entity, the Service Provider can decide to integrate the GW and WAN Edge PE functions in the same router for obvious CAPEX and OPEX saving reasons. This is illustrated in Figure 1. There are two interconnection solutions as follows:

1. End to end NVO3 tunnel across different data centers. NVE1 perform NVO3 encapsulation for DCI interconnection with NVE3, the destination VTEP IP is NVE3's IP. The GW doesn't perform NVO3 tunnel termination. The DCI WAN is pure underlay network.

2. Segmented NVO3 tunnels across different data centers. NVE1

doesn't perform end to end NVO3 encapsulation to NVE3 for DCI interconnection. The GW performs NVO3 tunnel encapsulation termination, and then transmits the inner original traffic through MPLS network to peer data center GW. The peer data center GW

terminates MPLS encapsulation, and then performs NVO3 encapsulation to transmit the traffic to local NVE3.

In the first solution, to differentiate bandwidth and QOS among different tenants or applications, different TE tunnels in the WAN network will be used to carry the end to end NVO3 encapsulation traffic using VN ID, NVO3 outer header DSCP and etc as traffic classification match part. BGP Flow-spec protocol can be used to set the traffic classification on all GWs simultaneously.

In the second solution, a centralized BGP speaker can be deployed for DDOS mitigation in the WAN network. When the analyzer detects abnormal traffic, it will automatically generate Flow-spec rules and distribute it to each GW through BGP Flow-spec protocol, the match part should include inner or outer L2/L3 layer or NVO3 header.

In summary, the Flow specification match part on the GW/PE should include inner layer 2 Ethernet header, inner layer 3 IP header, outer layer 2 Ethernet header, outer layer 3 IP header, and/or NVO3 header information. Because the current match part lacks layer indicator and NVO3 header information, so it can't be used directly for the traffic filtering based on NVO3 header or a specified layer header directly. This draft will propose a new subset of component types to support the NVO3 flow-spec application.

2. The Flow Specification encoding for NVO3

In default, the current flow-spec rules can only impose on the outer layer header of NVO3 encapsulation data packets. To make traffic filtering based on NVO3 header and inner header of NVO3 packets, a new component type acts as a delimiter is introduced. The delimiter type is used to specify the boundary of the inner or outer layer component types for NVO3 data packets. All the component types defined in [RFC5575],[IPv6-FlowSpec],[Layer2-FlowSpec],and etc can be used between two delimiters.

The NVO3 outer layer address normally belongs to public network, the

"Flow Specification" NLRI only for the outer layer header doesn't need to include Route Distinguisher field (8 bytes). If the outer layer address belongs to a VPN, the NLRI format for the outer header should consist of a fixed-length Route Distinguisher field (8 bytes) corresponding to the VPN, the RD is followed by the detail flow specifications for the outer layer.

VN ID is the identification for each tenant network, the "Flow Specification" NLRI for NVO3 header part should always include VN ID field, Route Distinguisher field doesn't need to be included.

The inner layer MAC/IP address always associates with a VN ID, the NLRI format for the inner header should consist of a fixed-length VNID field (4 bytes), the VNID is followed by the detail flow specifications for the inner layer. The NLRI length field shall include both the 4 bytes of the VN ID as well as the subsequent flow specification. In NVO3 terminating into VPN scenario, if multiple access VN ID maps to one VPN instance, one share VN ID can be carried in the Flow-Spec rule to enforce the rule to entire VPN instance, the share VN ID and VPN correspondence should be configured on each VPN PE beforehand, the function of the layer3 VN ID is same with Route Distinguisher to act as the identification of VPN instance.

This document proposes the following extended specifications for NVO3 flow:

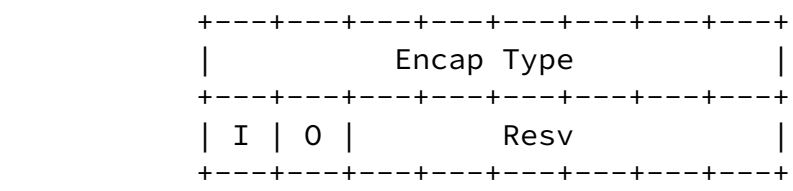
Type TBD1 - - Delimiter type

Encoding: <type (1 octet), length (1 octet), Value>.

When the delimiter type is present, it indicates the component types for the inner or outer layer of NVO3 packets will be followed immediately. At the same time, it indicates the end of the component types belonging to the former delimiter.

The value field defines encapsulation type and is encoded as:

0 1 2 3 4 5 6 7



This document defines the following Encap types:

- VXLAN: Tunnel Type = 0

- NVGRE: Tunnel Type = 1

I: If I is set to one, it indicates the component types for the inner layer of NVO3 packets will be followed immediately.

O: If O is set to one, it indicates the component types for the outer layer of NVO3 packets will be followed immediately.

For NVO3 header part, the following additional component types are introduced.

Type TBD2 - VNID

Encoding: <type (1 octet), [op, value]+>.

Defines a list of {operation, value} pairs used to match 24-bit VN ID which is used as tenant identification in NVO3 network. For NVGRE encapsulation, the VNID is equivalent to VSID. Values are encoded as 1- to 3-byte quantities.

Type TBD3 - Flow ID

Encoding: <type (1 octet), [op, value]+>

Defines a list of {operation, value} pairs used to match 8-bit Flow id fields which are only useful for NVGRE encapsulation. Values are encoded as 1-byte quantity.

3. The Flow Specification Traffic Actions for NVO3

The current traffic filtering actions can still be used for NVO3 encapsulation traffic. For Traffic Marking, only the DSCP in outer header can be modified.

4. Security Considerations

No new security issues are introduced to the BGP protocol by this specification.

5. IANA Considerations

IANA is requested to create and maintain a new registry entitled:

"Flow spec NVO3 Component Types":

Type TBD1 - Delimiter type

Type TBD2 - VNID

Type TBD3 - Flow ID

5.1. Normative References

[1] [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[2] [GENEVE] J. Gross, T. Sridhar, etc, " Geneve: Generic Network Virtualization Encapsulation", draft-ietf-nvo3-geneve-00, May 2015.

[3] [GUE] T. Herbert, L. Yong, O. Zia, " Generic UDP Encapsulation", draft-ietf-nvo3-gue-01, Jun 2015.

[4] [GPE] P. Quinn,etc, " Generic Protocol Extension for VXLAN", draft-ietf-nvo3-vxlan-gpe-00, May 2015.

5.2. Informative References

[1] [EVPN-Overlays] A. Sajassi,etc, " A Network Virtualization Overlay Solution using EVPN", draft-ietf-bess-evpn-overlay-01 , work in progress, February, 2014.

[2] [Inter-Overlays] J. Rabadan,etc, " Interconnect Solution for EVPN Overlay networks", draft-ietf-bess-dci-evpn-overlay-01, work in progress, July, 2015.

[3] [RFC7348] M. Mahalingam, etc, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC7348, August 2014.

[4] [NVGRE] P. Garg, etc, "NVGRE: Network Virtualization using Generic Routing Encapsulation", draft-sridharan-virtualization-nvgre-08, April 13, 2015.

[5] [IPv6-FlowSpec] R. Raszuk, etc, " Dissemination of Flow Specification Rules for IPv6", draft-ietf-idr-flow-spec-v6-06, November 2014.

[6] [Layer2-FlowSpec] W. Hao, etc, "Dissemination of Flow Specification Rules for L2 VPN", draft-ietf-idr-flowspec-l2vpn-02, August 2015.

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

6. Acknowledgments

The authors wish to acknowledge the important contributions of Jeff Haas, Susan Hares, Qiandeng Liang, Nan Wu, Yizhou Li, Lucy Yong.

Authors' Addresses

Weiguo Hao
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China
Email: haoweiguo@huawei.com

Shunwan Zhuang
Huawei Technologies

Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China
Email: zhuangshunwan@huawei.com

Zhenbin Li
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China
Email: lizhenbin@huawei.com

Rong Gu
China Mobile
gurong_cmcc@outlook.com