Idr Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 19, 2015

J. You
Q. Liang
Huawei
September 15, 2014

BGP VPN Peer Discovery Method
draft-you-idr-bgp-vpn-discovery-00

Abstract

This document proposes a VPN peer discovery method based on BGP. All BGP peers need to establish sessions with a centralized control point which collects and distributes BGP VPN Peer information according to local policies or filtering policies subscribed from BGP peers. Thereafter, the BGP node can select the interested peers with the same VPN services to establish sessions. This would avoid unnecessary session between uninterested BGP 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

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 March 19, 2015.

Copyright Notice

Table of Contents

1. Introduction

The base BGP-4 specification ([RFC4271]) utilizes TCP for session establishment between peers, which requires prior knowledge of the end point's address to which a BGP session should be targeted. The BGP auto discovery [I-D.raszuk-idr-bgp-auto-discovery] describes a method for automating portions of a router's BGP configuration via discovery of BGP peers with which to establish further sessions from an initial "bootstrap" router. However, in [I-D.raszuk-idr-bgp-auto-discovery], the VPN information of BGP peers is not collected by the "bootstrap" router. Instead, the VPN-related information is obtained after the BGP session established between BGP peers ([RFC4761], [RFC6624]). Basically, if there's no same VPN service between the BGP nodes, there's no need to establish the BGP session. If BGP node can obtain the VPN information of the BGP peers before BGP session establishment, it could avoid unnecessary session between uninterested BGP peers.

This document proposes a VPN peer discovery method based on BGP. All BGP peers need to establish sessions with a centralized control point which collects and distributes BGP VPN Peer information according to local policies or filtering policies subscribed from BGP peers.

Thereafter, the BGP node can select the interested peers with the same VPN services to establish sessions. From OAM aspects, it is beneficial for BGP node only following the interested BGP peers based on its filtering policies e.g. BGP peers in a specified AS. So this BGP node doesn't need to follow the VPN information of the BGP peers in other ASs. Thus, the BGP session between uninterested BGP peers will not be established.

2. Terminology

This section contains definitions for terms used frequently throughout this document. However, many additional definitions can be found in [RFC4760] and [RFC5575].

AFI: Address Family Identifier

AS: Autonomous System number

NLRI: Network Layer Reachability Information

RD: Route Distinguisher

RR: Route Reflector

SAFI: Subsequent Address Family Identifier

3. VPN Peer Discovery Method

The centralized control point collects all the VPN service information of BGP nodes within the domain, and then distributes the BGP VPN Peer information to the BGP peers according to local policies or filtering policies subscribed from BGP peers.

3.1. VPN Peer Filter

A new Path Attribute, i.e. VPN Peer Filter Attribute (Type Code:

TBD1) is defined to carry VPN peer filtering information. It is carried in the BGP Update message. This attribute is used by the BGP peer for subscribing the interested VPN peer information towards the centralized control point.

The format of VPN Peer Filter Attribute is defined in Figure 1:

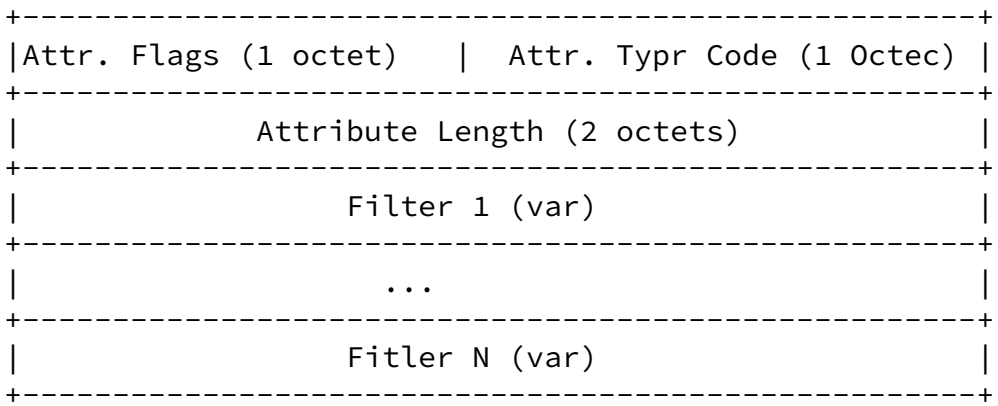

Figure 1: VPN Peer Filter Attribute

The attribute flags and type code fields are detailed in Figure 2:

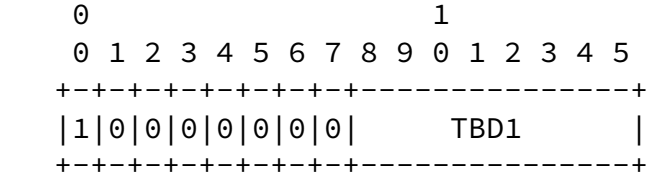

Figure 2: Flags & Type Code Fields

Bit 0: Optional attribute (value 1)

Bit 1: Non transitive attribute (value 0)

Bit 2: Partial bit (value 0 for optional non transitive

attributes)

Bit 3: Length of one octet (value 0)

Bit 4-7: Unused (value all zeros)

Type code: Attribute type code TBD1

Each VPN Peer Filter Attribute contains one or more filters. The Filter is encoded as: <type (1 octet), [op, value]+> ([RFC5575]). It contains a set of {operator, value} pairs that are used to match the specified value of the corresponding type defined in filter encoding. The operator byte is encoded as:

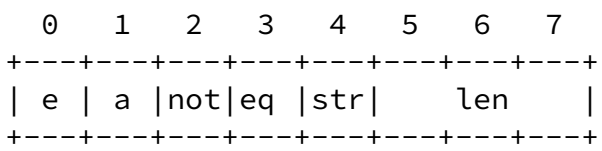

Figure 3: Numeric Operator

e: end-of-list bit. Set in the last {op, value} pair in the list.

a: AND bit. If unset, the previous term is logically ORed with the current one. If set, the operation is a logical AND. It should be unset in the first operator byte of a sequence. The AND operator has higher priority than OR for the purposes of evaluating logical expressions.

not: NOT bit. If set, logical negation of operation.

eq: equality between data and value.

str: STRING bit. If set, the value is an ASCII string.

len: The length of the value field for this operand is given as (1

<< len).

For the filter, the following types are defined:

type 1 - RD: RD, 8 octets

type 2 - AS: AS, 4 octets

type 3 - AFI/SAFI: AFI/SAFI, 3 octets

type 4 - Service Type: Service Type, 2 octets, to identify a VPN service type, e.g. EVPN, Kompella VPLS [RFC4761], BGP AD VPLS [RFC6074], IP VPN.

type 5 - Peer IP: Peer IP, 4 octets for IPv4 or 16 octets for IPv6.

type 6 - Export RT: RT/VPN target, 8 octets, filled with local import RTs/VPN targets of the subscriber.

3.2. VPN Peer Information

The BGP Speaker can use the Network Layer Reachability Information field of the MP_REACH_NLRI ([RFC4760]) attribute to notify the peer the corresponding VPN peer information, which is satisfied with local or peer's filtering policies. Similarly, The BGP Speaker can use the

Withdrawn Route NLRIs field of the MP_UNREACH_NLRI ([RFC4760]) attribute to notify the peer of removing the corresponding VPN peer information.

Besides, this document defines a new value for the Address Family Identifier field carried in the MP_REACH_NLRI and MP_UNREACH_NLRI attributes:

3 (TBD2) - Network Layer Reachability Information used for VPN Service

For the Subsequent Address Family Identifier field carried in the MP_REACH_NLRI and MP_UNREACH_NLRI attributes, if AFI = 3 (TBD), then the SAFI could be set to 0 by the sender, and it is ignored by the receiver.

The Network Layer Reachability information is encoded as one or more 2-tuples of the form <length, prefix> ([RFC4760]).

The extended Network Layer Reachability information field of the MP_REACH_NLRI or the Withdrawn Route NLRIs field of the MP_UNREACH_NLRI is encoded as shown in Figure 4, therein the VPN information including service type, RD, AS, AFI/SAFI and peer address can be regarded as a particular prefix.

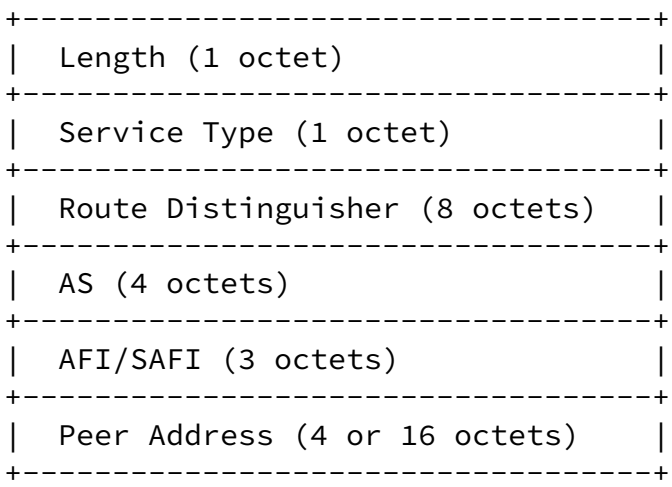

Figure 4: VPN Service Informatiom

The use and meaning of these fields are as follows:

Length: The Length field indicates the length of the VPN service information, including service type, RD, AS, AFI/SAFI and peer address.

Service Type: Service type is used to identify a VPN service type, e.g. EVPN, Kompella VPLS [RFC4761], BGP AD VPLS [RFC6074], IP VPN.

Route Distinguisher (RD): An 8-byte Route Distinguisher

AS: Autonomous System number

Address Family Identifier (AFI): This field in combination with the Subsequent Address Family Identifier field identifies the set of Network Layer protocols to which the address carried in the Next Hop field must belong, the way in which the address of the next hop is encoded, and the semantics of the Network Layer Reachability Information that follows. If the Next Hop is allowed to be from more than one Network Layer protocol, the encoding of the Next Hop MUST provide a way to determine its Network Layer protocol. Presently defined values for the Address Family Identifier field are specified in the IANA's Address Family Numbers registry.

Subsequent Address Family Identifier (SAFI): This field in combination with the Address Family Identifier field identifies the set of Network Layer protocols to which the address carried in the Next Hop must belong, the way in which the address of the next hop is encoded, and the semantics of the Network Layer Reachability Information that follows. If the Next Hop is allowed to be from more than one Network Layer protocol, the encoding of the Next Hop MUST provide a way to determine its Network Layer protocol.

Peer Address: the address of the peer, 4 bytes for IPv4, and 16 bytes for IPv6.

3.3. Deployment Considerations

The centralized control point can be deployed on the RR. Then the RR needs to establish the sessions with all the PEs to obtain the interested VPN peer information.

BGP nodes use VPN Peer Filter Attribute (Type Code: TBD1) for subscribing the interested VPN peer information towards the centralized control point.

The RR notifies the BGP peer the corresponding VPN peer information (an extended Network Layer Reachability Information), which is satisfied with local or peer's filtering policies.

4. IANA Considerations

This document defines a new Path Attribute, i.e. VPN Peer Filter Attribute (Type Code: TBD1), which will be used to carry VPN peer filtering information. A new attribute type code TBD1 is to be assigned by IANA from the BGP path attribute Type Code space.

This document defines a new MP_REACH_NLRI /MP_UNREACH_NLRI AFI type code TBD2 which will be used to carry VPN service. That value will need to be assigned by IANA from BGP AFI Type Code space.

This document defines a new NLRI format, to be carried in MP_REACH_NLRI and MP_UNREACH_NLRI attributes.

5. Security considerations

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

6. Acknowledgement

TBD.

7. References

7.1. Normative References

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

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[RFC4761] Kompella, K. and Y. Rekhter, "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, 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.

[RFC6074] Rosen, E., Davie, B., Radoaca, V., and W. Luo, "Provisioning, Auto-Discovery, and Signaling in Layer 2 Virtual Private Networks (L2VPNs)", RFC 6074, January 2011.

[RFC6624] Kompella, K., Kothari, B., and R. Cherukuri, "Layer 2 Virtual Private Networks Using BGP for Auto-Discovery and Signaling", RFC 6624, May 2012.

7.2. Informative References

[I-D.raszuk-idr-bgp-auto-discovery] Raszuk, R., Kumari, W., Mitchell, J., Patel, K., and J. Scudder, "BGP Auto Discovery", draft-raszuk-idr-bgp-auto-discovery-01 (work in progress), August 2014.

Authors' Addresses

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

Email: youjianjie@huawei.com

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

Email: liuweihang@huawei.com