Constrained Route Distribution for BGP based Virtual Private Networks (VPNs)
draft-dong-idr-vpn-route-constrain-00.txt

Status of this Memo

This Internet-Draft is submitted to IETF 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/ietf/1id-abstracts.txt.

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

This Internet-Draft will expire on August 15, 2009.

Copyright Notice

Copyright (c) 2009 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 in effect on the date of publication of this document (http://trustee.ietf.org/license-info). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.
Abstract

This document defines generalized procedures that allow BGP speakers to exchange Route Target reachability information. This information can be used to precisely control the propagation of different kinds of Virtual Private Network (VPN) routing information. This method avoids unnecessary advertising of VPN routes when more than one kind of VPN is deployed in the network.

Table of Contents

1. Introduction .................................................. 2
   1.1. Terminologies............................................ 3
2. Conventions used in this document ................................ 3
3. Multiple Kinds of VPNs in Service Provider Network ............ 3
   3.1. Multiple VPNs with Different AFIs........................ 4
      3.1.1. IPv4 L3VPN and IPv6 L3VPN.......................... 4
      3.1.2. L3VPN and L2VPN.................................... 5
   3.2. Multiple VPNs with Different SAFIs....................... 6
4. Considerations about Route Constraint in IPv6 L3VPN ............ 8
5. Considerations about Route Constraint in L1VPN .................. 8
6. Proposed Solutions............................................ 8
   6.1. Length of Route Target Field............................ 8
   6.2. Identifying RT Membership of Different VPNs............. 8
      6.2.1. Identifying AFI of VPNs............................ 8
      6.2.2. Identifying SAFI of VPNs........................... 9
      6.2.3. Proposed Format of RT Membership NLRI.............. 9
7. Security Considerations........................................ 10
8. IANA Considerations.......................................... 10
9. References.................................................... 10
   9.1. Normative References................................... 10
   9.2. Informative References................................. 11
Authors’ Addresses............................................. 12

1. Introduction

BGP [RFC4271] has been widely used in many kinds of Virtual Private Networks (VPNs) for exchanging routes and auto discovery information. Route Target (RT) extended communities defined in [RFC4360] are used to control the distribution of received information into VRFs.

[RFC4684] defines some procedures to restrict the distribution of VPN routes. It defines a new MP-BGP NLRI with [AFI=1, SAFI=132] for carrying RT membership information, which can be used to control the propagation of VPN routes. The procedures are helpful in limiting the propagation of VPN routes in networks where only one kind of VPN is...
deployed. If multiple different kinds of VPNs are used in the network, the procedures in RFC 4684 are not sufficient for providing precise control of VPN route distribution.

This document analyses possible problems RFC 4684 may meet with and also provides solutions for these problems.

1.1. Terminologies

Terms and acronyms specific to BGP and VPNs are listed below:

- AFI: Address Family Identifier
- CE: Customer Edge
- L2VPN: Layer 2 Virtual Private Network
- L3VPN: Layer 3 Virtual Private Network
- MVPN: Multicast L3VPN
- NLRI: Network Layer Reachability Information
- PE: Provider Edge
- RT: Route Target
- SAFI: Subsequent Address Family Identifier
- VPLS: Virtual Private LAN Service
- VPN: Virtual Private Network

2. Conventions used in this document

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].

3. Multiple Kinds of VPNs in Service Provider Network

In many scenarios, it is possible that Service Provider (SP) needs to deploy more than one kind of VPNs in their network. For example, one SP may deploy both IPv4 L3VPN and IPv6 L3VPN, or both L3VPN and L2VPN, or both unicast VPN and multicast VPN, etc. Note that using the procedures of RFC 4684, the Route Target (RT) membership NLRI used for controlling distribution of VPNs routes cannot tell which kind of
VPN routes are wanted. Thus in these scenarios, the procedures of RFC 4684 are not sufficient to provide precise control for VPN route distribution, some PEs may receive unwanted VPN routes. Detailed analyses are given in sections below.

3.1. Multiple VPNs with Different AFIs

Service providers may deploy multiple VPNs with different Address Family Identifiers (AFI), such as IPv4 L3VPN (AFI=1), IPv6 L3VPN (AFI=2) and L2VPN (AFI=25). When more than one of these kinds of VPNs is deployed in the same network, using procedures of RFC 4684 can lead to unwanted routes being received by some PEs.

3.1.1. IPv4 L3VPN and IPv6 L3VPN

Some service providers need to deploy both IPv4 L3VPN (VPNv4) and IPv6 L3VPN (VPNv6) in their network. The RTs used for VPNv6 can be the same as the ones for VPNv4. However, it’s likely that not all the VPN sites are both IPv4 and IPv6 capable, some of them may be only IPv4 capable, some other ones may support both IPv4 and IPv6, and the others can only recognize IPv6 packets.

In Figure 1, suppose the VPN site of VPN-1 connected to PE-1 is only IPv4 capable, the sites of VPN-1 connected to PE-2 and PE-3 are IPv6 capable. Using procedures of RFC 4684, PE-1 advertises the RT membership information of VPN-1 to the other PEs. According to the received RT membership information, PE-2 and PE-3 will advertise VPNv6 routes of VPN-1 to PE-1. As a result, PE-1 will receive unwanted VPNv6 routes of VPN-1. Similarly, PE-1 can also receive unwanted VPNv4 routes of VPN-2.
Even if RTs allocated for VPNv6 are different from the ones used for VPNv4 on each PE, there can be some overlapping between the RT space of VPNv4 and VPNv6 in SP backbone network, in which case some PEs can also receive unwanted VPN routes of other VPNs.

### 3.1.2. L3VPN and L2VPN

The mechanisms defined in RFC 4684 are claimed to be applicable for L2VPNs which use Route Targets to control distribution of routing information. However, if both L3VPN and L2VPN are deployed in the same network, the mechanisms defined in RFC 4684 are not sufficient for realizing accurate control of route distribution.

If there is some overlapping between the RT space of L3VPNs and L2VPNs in the network, PEs will receive unwanted VPN routes of another kind of VPN.

For example, in Figure 2, RT-1 is used by PE-1 and PE-4 for L2VPN-1, and is also used by PE-2 and PE-3 for L3VPN-3. If PE-1 and PE-4 advertise RT membership information of RT-1 to other PEs in the
network, subsequently PE-2 and PE-3 would advertise unwanted L3VPN routes to PE-1 and PE-4.

![Diagram of network showing VPN-1, VPN-2, VPN-3 with PE-1, PE-2, PE-3, PE-4, RT-1, RT-2, L2VPN, L3VPN]

Figure 2.

### 3.2. Multiple VPNs with Different SAFIs

When the AFI value is the same, different kinds of VPN routes are further differentiated using Subsequent Address Family Identifiers (SAFI). Such routes may be needed by different sets of PEs. However, the procedures of [RFC 4684](http://rfc-editor.org/rfc/rfc4684.txt) cannot describe the requirements of routes with a particular SAFI.

For example, Multicast L3VPN [MVPN, MVPN-BGP] and VPLS multicast [VPLS-MCAST] have defined new SAFIs for exchanging multicast routing information, and Route Targets are also used in these scenarios to control distribution of multicast routing information.

Since MVPN is deployed on the basis of unicast VPN, they are always coexistent in the same network.
As [MVPN-BGP] says, by default the distribution of Intra-AS I-PMSI A-D route is controlled by the same Route Targets as the ones used for the distribution of VPN-IP unicast routes. Thus PE advertising RT membership NLRI may receive unwanted routing information, i.e., PE wants to receive only unicast VPN routes corresponding to the RT may also receive unwanted multicast routing information.

In Figure 3, suppose the site of VPN-1 connected to PE-1 only support IP unicast, the sites of VPN-1 connected to PE-2 and PE-3 support IP multicast. Using the mechanisms defined in [RFC 4684], PE-1 will advertise the RT membership information of VPN-1 to the other PEs. According to the received RT membership information, PE-2 and PE-3 will advertise multicast routes of VPN-1 to PE-1. As a result, PE-1 will receive unwanted multicast routes of VPN-1.

Even if RTs allocated for MVPNs are different from the ones used for unicast VPNs on each PE, there can be some overlapping between the RT space of MVPNs and unicast VPNs in the network, in which case some PEs can also receive unwanted VPN routes of other VPNs.
4. Considerations about Route Constraint in IPv6 L3VPN

The format of RT membership NLRI in RFC 4684 contains a Route Target field of 8 octets. However, [IPv6-EXT-COMM] has defined IPv6 address specific Route Target with the length of 20 octets, thus the format of RT membership NLRI is not applicable when IPv6 address specific Route Target is used. From this point of view, an update to the format of RT-membership NLRI is necessary.

5. Considerations about Route Constraint in L1VPN

BGP is also used for L1VPN auto discovery as described in [RFC5195]. The SAFI value 69 has been assigned for L1VPN auto discovery information. If route constrain procedures are used in L1VPN, then deploying both L1VPN and other kinds of VPNs in the same network can lead to problems similar to the examples in previous sections.

6. Proposed Solutions

This document proposes to extend RFC 4684 to a more general method for controlling the route distribution of all kinds of BGP based VPNs in any scenario.

6.1. Length of Route Target Field

First of all, this document proposes to change the length of Route Target field in RT membership NLRI to "variable". Thus the NLRI format is compatible with IPv6 address specific Route Target and other new types of Route Targets which can be defined in future. Since the RT membership NLRI is encoded as defined in section 4 of RFC 2858 (which is obsoleted by [RFC4760]), thus the length field in RT membership NLRI can be used to calculate the length of Route Target.

6.2. Identifying RT Membership of Different VPNs

6.2.1. Identifying AFI of VPNs

In order to identify corresponding AFI of VPN routes that the RT membership NLRI stands for, this document proposes to extend the AFI value of RT membership NLRI. The AFI of this NLRI SHOULD be one of the AFIs that use Route Target to control route distribution. In order to be compatible with RFC 4684, this document defines a new SAFI called Generalized Route Target Membership. A new SAFI value 135 needs to be assigned by IANA. Thus the [AFI, SAFI] value pair of the Generalized RT Membership NLRI could be [AFI=1, SAFI=135] for IPv4 L3VPN, and [AFI=2, SAFI=135] for IPv6 L3VPN, and [AFI=25, SAFI=135]...
for L2VPN, etc. Since currently there is no fixed AFI value for L1VPN, a new AFI value MAY need to be allocated by IANA, and the [AFI=TBD, SAFI=135] value pair represents the Generalized RT membership NLRI of L1VPN.

6.2.2. Identifying SAFI of VPNs

In order to distinguish the control of different types of VPN routes with the same AFI value, e.g. unicast L3VPN and MVPN, or unicast L2VPN and multicast L2VPN, the Generalized RT membership NLRI MUST contain information about the SAFI value of the VPN routes being requested. Thus the Generalized RT membership NLRI can be accurately identified as RT information of one particular type of VPN route. Besides, if L1VPN and other kinds of VPNs are deployed in the same network, and no fixed AFI value has been allocated for L1VPN, the SAFI value of L1VPN is needed to identify RT information of L1VPN.

6.2.3. Proposed Format of Generalized RT Membership NLRI

The format of Generalized RT membership NLRI is structured as follows:

```
+-------------------------------+
| Length            (1 octet)   |
+-------------------------------+
| Origin AS        (4 octets)   |
+-------------------------------+
| SAFI of VPN       (1 octet)   |
+-------------------------------+
|                               |
|                               |
| ~ Route Target     (variable) ~|
+-------------------------------+
```

The Length field is used to identify the total length of the rest fields. [Author notes: Though this field is defined as "the length in bits" in [RFC4760], it is RECOMMENDED that it represents the length in octets of the rest fields as in [RFC4761] for convenience.]

The Origin AS field contains an Autonomous System number. Two octets AS numbers are encoded in the two low order octets of the Origin AS field, with the two high order octets set to zero.

The "SAFI of VPN" field is the SAFI value of the VPN routes the PE wants to import using the Route Target below.
The Route Target field contains Route Target of VPN routes being requested. Note the length of the Route Target field is variable, and can be calculated using the Length field of this NLRI.

7. Capability advertisement

In order for two BGP speakers to exchange Generalized RT membership NLRI, they MUST use BGP Capabilities Advertisement to ensure that they both are capable of properly processing such NLRI. This is done as specified in RFC4760, by using capability code 1 (multiprotocol BGP) with an AFI of the corresponding VPN and an SAFI of 135.

8. Security Considerations

This document does not change the security properties of BGP based VPNs and RFC 4684.

9. IANA Considerations

IANA needs to assign the SAFI value 135 for Generalized Route Target Membership. This code point will come from the "Subsequent Address Family Identifiers" registry.

IANA MAY assign an AFI number for L1VPN. This code point will come from the "Address Family Numbers" registry.

10. References

10.1. Normative References


10.2. Informative References


Authors’ Addresses

Jie Dong
Huawei Technologies Co., Ltd
KuiKe Building, No.9 Xinxi Rd.,
Hai-Dian District
Beijing, 100085
P.R. China
EMail: dongjie_dj@huawei.com

Mach (Guoyi) Chen
Huawei Technologies Co., Ltd
KuiKe Building, No.9 Xinxi Rd.,
Hai-Dian District
Beijing, 100085
P.R. China
EMail: mach@huawei.com

Guiyan Liu
Huawei Technologies Co., Ltd
Huawei Building, No.156 Beiqing Rd.,
Hai-Dian District
Beijing, 100095
P.R. China
EMail: 162547@huawei.com

Hui Ni
Huawei Technologies Co., Ltd
Huawei Building, No.156 Beiqing Rd.,
Hai-Dian District
Beijing, 100095
P.R. China
EMail: nihui@huawei.com