Network Working Group	X. Xu	
Internet-Draft	Huawei	
Intended status: Informational	P. Francis	
Expires: April 3, 2010	MPI-SWS	
	September 30, 2009	

TOC

Proposal to use an inner MPLS label to identify the remote ASBR VA
draft-ietf-grow-va-mpls-innerlabel-00.txt

Status of this Memo

Copyright Notice

Abstract

The draft "MPLS Tunnels for Virtual Aggregation" [I-D.ietf-grow-va-mpls] (Francis, P. and X. Xu, “MPLS Tunnels for Virtual Aggregation,” May 2009.) specifies how MPLS is used as the tunneling protocol for Virtual Aggregation (VA). The -00 version of that draft specifies only one level of labels, with the result that one

Label Switched Path (LSP) for every remote ASBR must be established. For large ISPs, this can amount to a large number of LSPs. This draft proposes adding the option of using an inner label to identify the remote ASBR. Either an outer label or an IP tunnel is used to reach the local ASBR. When MPLS is used as the tunneling protocol, this reduces the number of LSPs to the number of local border routers (ASBR).

Table of Contents

1. Proposal
    1.1. Requirements notation
    1.2. Changes from Previous Versions
2. IANA Considerations
3. Security Considerations
4. Normative References
§ Authors' Addresses

---

1. Proposal

TOC

The draft "MPLS Tunnels for Virtual Aggregation" [I-D.ietf-grow-va-mpls] (Francis, P. and X. Xu, “MPLS Tunnels for Virtual Aggregation,” May 2009.) specified how MPLS is used as the tunneling protocol for Virtual Aggregation (VA). The -00 version of that draft specifies only one level of labels, with the result that one LSP for every remote ASBR must be established. For large ISPs, this can amount to a large number of LSPs (roughly 20,000 for one large ISP we studied). This draft proposes the optional use of an inner label to reduce the number of LSPs to the number of local ASBRs. Besides improving the efficiency of VA, this also makes it feasible to use MPLS TE (traffic engineered) LSPs.

VA requires that tunneled packets are "targeted" to remote ASBRs. However, the tunnel header must be stripped before the packet is transmitted to the remote ASBR. This means that the tunnel header must identify the remote ASBR to the local ASBR, so that the local ASBR may strip the header and forward the packet to the remote ASBR. In the -00 draft of [I-D.ietf-grow-va-mpls] (Francis, P. and X. Xu, “MPLS Tunnels for Virtual Aggregation,” May 2009.), there is one LSP per remote ASBR. In other words, there is a distinct label per remote ASBR.

This draft proposes adding the option of using an inner label to identify the remote ASBR. Either an outer label or an IP tunnel identifies the local ASBR. When the local ASBR receives the packet, it strips off the outer label/header, uses the value of the inner label to identify the remote ASBR, and then strips the inner label before forwarding the packet to the remote ASBR. Note that, in the case of

stacked labels, the outer label may have been stripped by the previous hop using penultimate hop popping (PHP).
This style of tunneling is essentially identical to that used for MPLS VPNs [RFC4364] (Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)," February 2006.), though simpler because there is no need for virtual forwarding tables.
There are three forms of tunneling that can be used, stacked labels ([RFC3032] (Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding," January 2001.)), and MPLS-in-IP or MPLS-in-GRE ([RFC4023] (Worster, T., Rekhter, Y., and E. Rosen, "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)," March 2005.)), as follows:

```
Stacked labels (RFC3032):
  Payload | IP | Inner label | Outer label | link | ==>

MPLS-in-IP (RFC4023):
  Payload | IP | Inner label | Outer IP header | link | ==>

MPLS-in-GRE (RFC4023):
  Payload | IP | Inner label | GRE | Outer IP header | link | ==>
```

When a local ASBR advertises a route into iBGP, it sets the Next Hop to itself, and assigns a label to the route. This label is used as the inner label, and identifies the remote ASBR from which the route was received [RFC3107] (Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4," May 2001.).
The presence of the inner label in the iBGP update acts as the signal to the receiving router that an inner label should be used in packets tunneled to the Next Hop address. Other information is used to determine whether the tunnel itself is MPLS, IP, or GRE. Specifically, [I-D.ietf-grow-va-gre] (Francis, P., Raszuk, R., and X. Xu, "GRE and IP-in-IP Tunnels for Virtual Aggregation," July 2009.). specifies how to convey the use of IP or GRE tunneling in BGP for VA (i.e. though the attributes from [RFC5512] (Mohapatra, P. and E. Rosen, "BGP Encapsulation SAFI and BGP Tunnel Encapsulation Attribute," April 2009.)). If these attributes indicate IP or GRE tunneling, then the corresponding IP or GRE tunnel should be used. If no 5512 attribute is present, but there is a LSP to the Next Hop address, then the LSP should be used. If no 5512 attribute is present, and there is no LSP to the Next Hop address, then the packet should be IP tunneled to the Next Hop address.
The following table summarizes the tunneling behavior (and for completeness includes the both the cases where the inner label is and is not signaled).

Inner label?	5512 attr?	LSP to Next Hop?	Tunnel Behavior
No	No	No	Don't tunnel packet (normal behavior without VA)
No	No	Yes	Use LSP
No	Yes	No	Use 5512 tunnel to next hop
No	Yes	Yes	Use 5512 tunnel to Next Hop if possible, else use LSP *
Yes	No	No	Use IP tunnel to Next Hop with inner label
Yes	No	Yes	Use LSP (stacked labels)
Yes	Yes	No	Use 5512 tunnel to Next Hop with inner label
Yes	Yes	Yes	Use 5512 tunnel to Next Hop with inner label if possible, else use LSP *

* If the receiving router does not have the appropriate 5512 tunneling capability (IP or GRE), and it does have LSP capability, then it should use the LSP.

It is important to note that conveying inner label or tunneling information in BGP is not a negotiation per se: there is no assurance that the recipient of the information can actually do the type of tunneling indicated. It is therefore necessary for the AS administrator to insure that routers are capable of acting on any labeling or tunneling information that they receives.

1.1. Requirements notation

TOC

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] (Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” March 1997.).

1.2. Changes from Previous Versions

TOC

This is the first version of this draft.

---

2. IANA Considerations

TOC

There are no IANA considerations.

---

3. Security Considerations

TOC

Because this document describes a standard application of MPLS, there are no new security considerations beyond those already described in [I-D.ietf-grow-va-mpls] (Francis, P. and X. Xu, “MPLS Tunnels for Virtual Aggregation,” May 2009.). It is worth noting, however, that the some of the security considerations normally associated with VPNs, namely that it not be possible for a non-VPN source to inject a packet into a VPN, do not apply here. Virtual Aggregation applies to global routing, not to VPN, and therefore it is not necessary to isolate communities.

---

4. Normative References

TOC

[I-D.ietf-grow-va-gre]	Francis, P., Raszuk, R., and X. Xu, “GRE and IP-in-IP Tunnels for Virtual Aggregation,” draft-ietf-grow-va-gre-00 (work in progress), July 2009 (TXT).
[I-D.ietf-grow-va-mpls]	Francis, P. and X. Xu, “MPLS Tunnels for Virtual Aggregation,” draft-ietf-grow-va-mpls-00 (work in progress), May 2009 (TXT).
[RFC2119]	Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC3032]	Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, “MPLS Label Stack Encoding,” RFC 3032, January 2001 (TXT).
[RFC3107]	Rekhter, Y. and E. Rosen, “Carrying Label Information in BGP-4,” RFC 3107, May 2001 (TXT).
[RFC4023]	Worster, T., Rekhter, Y., and E. Rosen, “Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE),” RFC 4023, March 2005 (TXT).
[RFC4364]	Rosen, E. and Y. Rekhter, “BGP/MPLS IP Virtual Private Networks (VPNs),” RFC 4364, February 2006 (TXT).
[RFC5512]	Mohapatra, P. and E. Rosen, “BGP Encapsulation SAFI and BGP Tunnel Encapsulation Attribute,” RFC 5512, April 2009 (TXT).

Authors' Addresses

TOC

	Xiaohu Xu
	Huawei Technologies
	No.3 Xinxi Rd., Shang-Di Information Industry Base, Hai-Dian District
	Beijing, Beijing 100085
	P.R.China
Phone:	+86 10 82836073
Email:	xuxh@huawei.com
	Paul Francis
	Max Planck Institute for Software Systems
	Gottlieb-Daimler-Strasse
	Kaiserslautern 67633
	Germany
Email:	francis@mpi-sws.org