Network Working Group
Request for Comments: 4797
Category: Informational

Y. Rekhter
R. Bonica
Juniper Networks
E. Rosen
Cisco Systems, Inc.

January 2007

Use of Provider Edge to Provider Edge (PE-PE) Generic Routing Encapsulation (GRE) or IP in BGP/MPLS IP Virtual Private Networks

Status of This Memo

Copyright Notice

IESG Note

This document proposes an automated mechanism for establishing tunnels between provider-edge routers in a VPN, but does not provide an automated mechanism for establishing security associations for these tunnels. Without such a mechanism, this document is not appropriate for publication on the Internet standards track.

Abstract

This document describes an implementation strategy for BGP/MPLS IP Virtual Private Networks (VPNs) in which the outermost MPLS label (i.e., the tunnel label) is replaced with either an IP header or an IP header with Generic Routing Encapsulation (GRE).

The implementation strategy described herein enables the deployment of BGP/MPLS IP VPN technology in networks whose edge devices are MPLS and VPN aware, but whose interior devices are not.

Table of Contents

1. Introduction

A "conventional" BGP/MPLS IP VPN [2] is characterized as follows:

- Each Provider Edge (PE) router maintains one or more Virtual Routing and Forwarding (VRF) instances. A VRF instances is a VPN-specific forwarding table.

- PE routers exchange reachability information with one another using BGP [3] with multi-protocol extensions [4].

- MPLS Label Switching Paths (LSPs) [5] connect PE routers to one another.

In simple configurations, the VPN service is offered by a single Autonomous System (AS). All service provider routers are contained by a single AS and all VPN sites attach to that AS. When an ingress PE router receives a packet from a VPN site, it looks up the packet's destination IP address in a VRF that is associated with the packet's ingress attachment circuit. As a result of this lookup, the ingress PE router determines an MPLS label stack, a data link header, and an output interface. The label stack is prepended to the packet, the data link header is prepended to that, and the resulting frame is queued for the output interface.

The innermost label in the MPLS label stack is called the VPN route label. The VPN route label is significant and visible to the egress PE router only. It controls forwarding of the packet by the egress PE router.

The outermost label in the MPLS label stack is called the tunnel label. The tunnel label causes the packet to be delivered to the egress PE router that understands the VPN route label. Specifically, the tunnel label identifies an MPLS LSP that connects the ingress PE router to the egress PE router. In the context of BGP/MPLS IP VPNs, this LSP is called a tunnel LSP.

The tunnel LSP provides a forwarding path between the ingress and egress PE routers. Quality of service (QoS) information can be mapped from the VPN packet to the tunnel LSP header so that required forwarding behaviors can be maintained at each hop along the forwarding path.

Sections 9 and 10 of reference [2] define more complex configurations (i.e., carriers' carrier and multi-AS backbones) in which service providers offer L3VPN services across multiple autonomous systems. In these configurations, VPN route labels can be stitched together

across AS boundaries. Within each AS, tunnel LSPs carry VPN packets from network edge to network edge.

In most configurations, tunnel LSPs never traverse AS boundaries. The tunnel LSP is always contained within a single AS. In one particular configuration (i.e., Inter-provider Option C), tunnel LSPs may traverse AS boundaries.

This memo describes procedures for creating an MPLS packet that carries the VPN route label, but does not carry the tunnel label. Then, using either GRE or IP encapsulation, the ingress PE router sends the MPLS packet across the network to the egress PE router.

That is, a GRE or IP tunnel replaces the tunnel LSP that was present in "conventional" BGP/MPLS IP VPNs. Like the tunnel LSP, the GRE/IP tunnel provides a forwarding path between the ingress and egress PE routers. QoS information can be copied from the VPN packet to the GRE/IP tunnel header so that required forwarding behaviors can be maintained at each hop along the forwarding path. However, because the GRE/IP tunnel lacks traffic engineering capabilities, any traffic engineering features provided by the tunnel LSP are lost.

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 RFC 2119 [1].

3. Motivation

"Conventional" BGP/MPLS IP VPNs require an MPLS Label Switched Path (LSP) between a packet’s ingress PE router and its egress PE router. This means that a BGP/MPLS IP VPN cannot be implemented if there is a part of the path between the ingress and egress PE routers that does not support MPLS.

In order to enable BGP/MPLS IP VPNs to be deployed even when there are non-MPLS routers along the path between the ingress and egress PE routers, it is desirable to have an alternative, which allows the tunnel label to be replaced with either an IP or (IP + GRE) header. The encapsulation header would have the address of the egress PE router in its destination IP address field, and this would cause the packet to be delivered to the egress PE router.

In this procedure, the ingress and egress PE routers themselves must support MPLS, but that is not an issue, as those routers must necessarily have BGP/MPLS IP VPN support, whereas the transit routers need not support MPLS or BGP/MPLS VPNs.

4. Specification

In short, the technical approach specified here is:

1. Continue to use MPLS to identify a VPN route, by continuing to add an MPLS label stack to the VPN packets. Between the ingress and egress PE router, the outermost member of the label stack will represent the VPN route label.

2. An MPLS-in-GRE or MPLS-in-IP [6] encapsulation will be used to turn the MPLS packet, described above, back into an IP packet. This, in effect, creates a GRE or an IP tunnel between the ingress PE router and the egress PE router.

The net effect is that an MPLS packet gets sent through a GRE or an IP tunnel.

Service providers must protect the above-mentioned IP or GRE tunnel as recommended in Section 8.2 of reference [6]. As stated in that document:

> "If the tunnel lies entirely within a single administrative domain, address filtering at the boundaries can be used to ensure that no packet with the IP source address of a tunnel endpoint or with the IP destination address of a tunnel endpoint can enter the domain from outside.
>
> However, when the tunnel head and the tunnel tail are not in the same administrative domain, this may become difficult, and filtering based on the destination address can even become impossible if the packets must traverse the public Internet.
>
> Sometimes only source address filtering (but not destination address filtering) is done at the boundaries of an administrative domain. If this is the case, the filtering does not provide effective protection at all unless the decapsulator of an MPLS-in-IP or MPLS-in-GRE validates the IP source address of the packet. This document does not require that the decapsulator validate the IP source address of the tunneled packets, but it should be understood that failure to do so presupposes that there is effective destination-based (or a combination of source-based and destination-based) filtering at the boundaries."

4.1. MPLS-in-IP/MPLS-in-GRE Encapsulation by Ingress PE

The following description is not meant to specify an implementation strategy; any implementation procedure that produces the same result is acceptable.

When an ingress PE router receives a packet from a Customer Edge (CE) router, it looks up the packet's destination IP address in a VRF that is associated with the packet's ingress attachment circuit. This enables the (ingress) PE router to find a VPN-IP route. The VPN-IP route will have an associated VPN route label and an associated BGP Next Hop. The label is pushed on the packet. Then an IP (or IP+GRE) encapsulation header is prepended to the packet, creating an MPLS-in-IP (or MPLS-in-GRE) encapsulated packet. The IP source address field of the encapsulation header will be an address of the ingress PE router itself. The IP destination address field of the encapsulation header will contain the value of the associated BGP Next Hop; this will be an IP address of the egress PE router. QoS information can be copied from the VPN packet to the GRE/IP tunnel header so that required forwarding behaviors can be maintained at each hop along the forwarding path.

The IP address of the remote tunnel endpoints MAY be inferred from the Network Address of the Next Hop field of the MP_REACH_NLRI BGP attribute [4]. Note that the set of Next Hop Network Addresses is not known in advance, but is learned dynamically via the BGP distribution of VPN-IP routes. Assuming a consistent set of tunnel capabilities exist between all the PEs and Autonomous System Border Routers (ASBRs), no a priori configuration of the remote tunnel endpoints is needed. The entire set of PE and ASBRs MUST have the same tunnel capabilities if the dynamic creation of IP (or GRE) tunnels is desired. The preference to use an IP (or GRE) tunnel MUST be configured. A set of PEs using two or more tunnel mechanisms (i.e., LSP, GRE, IP, etc.) MUST determine the tunnel type on a per-peer basis. The automatic association of tunnel capabilities on a per-peer basis is for future study. Note that these tunnels SHOULD NOT be IGP-visible links, and routing adjacencies SHOULD NOT be supported across these tunnel.

4.2. MPLS-in-IP/MPLS-in-GRE Decapsulation by Egress PE

Every egress PE is also an ingress PE, and hence has the ability to decapsulate MPLS-in-IP (or MPLS-in-GRE) packets. After decapsulation, the packets SHOULD be delivered to the routing function for ordinary MPLS switching.

As stated above, if the service provider deploys source-based filtering at network edges to protect the IP/GRE tunnel (instead of destination-based filtering), the decapsulator must validate the IP source address of the tunneled packets.

5. Implications on Packet Spoofing

It should be noted that if the tunnel MPLS labels are replaced with an unsecured IP encapsulation, like GRE or IP, it becomes more difficult to protect the VPNs against spoofed packets. This is because a Service Provider (SP) can protect against spoofed MPLS packets by the simple expedient of not accepting MPLS packets from outside its own boundaries (or more generally, by keeping track of which labels are validly received over which interfaces, and discarding packets that arrive with labels that are not valid for their incoming interfaces).

By contrast, protection against spoofed IP packets requires all SP boundary routers to perform filtering; either (a) filtering packets from "outside" the SP, which are addressed to PE routers, or (b) filtering packets from "outside" the SP, which have source addresses that belong "inside" and, in addition, filtering on each PE all packets that have source addresses that belong "outside" the SP.

The maintenance of these filter lists can be management intensive. Furthermore, depending upon implementation, these filter lists can be performance affecting. However, such filters may be required for reasons other than protection against spoofed VPN packets, in which case the additional maintenance overhead of these filters to protect (among other things) against spoofing of VPN packets may be of no practical significance. Note that allocating IP addresses used for GRE or IP tunnels out of a single (or a small number of) IP block could simplify maintenance of the filters.

6. Security Considerations

Security considerations in reference [6] apply here as well. Additional security issues are discussed in the previous section "Implications on Packet Spoofing".

7. Acknowledgments

Thanks to Robert Raszuk and Scott Wainner for their contributions to this document.

8. Normative References

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

[2] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

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

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

[5] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[6] Worster, T., Rekhter, Y., and E. Rosen, "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)", RFC 4023, March 2005.

Authors' Addresses

Yakov Rekhter
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

EMail: yakov@juniper.net

Ronald P. Bonica
Juniper Networks
2251 Corporate Park Drive
Herndon, VA 20171
US

EMail: rbonica@juniper.net

Eric C. Rosen
Cisco Systems, Inc.
1414 Massachusetts Avenue
Boxborough, MA 01719
US

EMail: erosen@cisco.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.