Internet Engineering Task Force (IETF)
Request for Comments: 6178
Updates: 3031
Category: Standards Track
ISSN: 2070-1721

D. Smith
J. Mullooly
Cisco Systems
W. Jaeger
AT&T
T. Scholl
nLayer Communications

March 2011

Label Edge Router Forwarding of IPv4 Option Packets

Abstract

This document specifies how Label Edge Routers (LERs) should behave when determining whether to MPLS encapsulate an IPv4 packet with header options. Lack of a formal standard has resulted in different LER forwarding behaviors for IPv4 packets with header options despite being associated with a prefix-based Forwarding Equivalence Class (FEC). IPv4 option packets that belong to a prefix-based FEC, yet are forwarded into an IPv4/MPLS network without being MPLS-encapsulated, present a security risk against the MPLS infrastructure. Further, LERs that are unable to MPLS encapsulate IPv4 packets with header options cannot operate in certain MPLS environments. While this newly defined LER behavior is mandatory to implement, it is optional to invoke.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6178.

Table of Contents

1. Motivation

This document is motivated by the need to formalize MPLS encapsulation processing of IPv4 packets with header options in order to mitigate the existing risks of IPv4 options-based security attacks against MPLS infrastructures. We believe that this document adds details that have not been fully addressed in [RFC3031] and [RFC3032], and that the methods presented in this document update [RFC3031] as well as complement [RFC3270], [RFC3443], and [RFC4950].

2. Introduction

The IPv4 packet header provides for various IPv4 options as originally specified in [RFC791]. IPv4 header options are used to enable control functions within the IPv4 data forwarding plane that are required in some specific situations but not necessary for most common IPv4 communications. Typical IPv4 header options include

provisions for timestamps, security, and special routing. Example IPv4 header options and applications include but are not limited to the following:

- o Strict and Loose Source Route Options: Used to IPv4 route the IPv4 packet based on information supplied by the source.
- o Record Route Option: Used to trace the route an IPv4 packet takes.
- o Router Alert Option: Indicates to downstream IPv4 routers to examine these IPv4 packets more closely.

The list of current IPv4 header options can be accessed at [IANA].

IPv4 packets may or may not use IPv4 header options (they are optional), but IPv4 header option handling mechanisms must be implemented by all IPv4 protocol stacks (hosts and routers). Each IPv4 header option has distinct header fields and lengths. IPv4 options extend the IPv4 packet header length beyond the minimum of 20 octets. As a result, IPv4 packets received with header options are typically handled as exceptions and in a less efficient manner due to their variable length and complex processing requirements. For example, many router implementations punt such IPv4 option packets from the hardware forwarding (fast) path into the software forwarding (slow) path causing high CPU utilization. Even when the forwarding plane can parse a variable-length header, it may still need to punt to the control plane because the forwarding plane may not have the clock cycles or intelligence required to process the header option.

Multi-Protocol Label Switching (MPLS) [RFC3031] is a technology in which packets associated with a prefix-based Forwarding Equivalence Class (FEC) are encapsulated with a label stack and then switched along a label switched path (LSP) by a sequence of label switch routers (LSRs). These intermediate LSRs do not generally perform any processing of the IPv4 header as packets are forwarded. (There are some exceptions to this rule, such as ICMP processing and LSP ping, as described in [RFC3032] and [RFC4379], respectively.) Many MPLS deployments rely on LSRs to provide layer 3 transparency much like ATM switches are transparent at layer 2. Such deployments often minimize the IPv4 routing information (e.g., no BGP transit routes) carried by LSRs since it is not necessary for MPLS forwarding of transit packets.

Even though MPLS encapsulation seems to offer a viable solution to provide layer 3 transparency, there is currently no formal standard for MPLS encapsulation of IPv4 packets with header options that belong to a prefix-based FEC. Lack of a formal standard has resulted

in inconsistent forwarding behaviors by ingress Label Edge Routers (LERs). When IPv4 packets are MPLS encapsulated by an ingress LER, for example, the IPv4 header including option fields of transit packets are not acted upon by downstream LSRs that forward based on the MPLS label(s). Conversely, when a packet is IPv4 forwarded by an ingress LER two undesirable behaviors can result. First, a downstream LSR may not have sufficient IPv4 routing information to forward the packet resulting in packet loss. Second, downstream LSRs must apply IPv4 forwarding rules that may expose them to IPv4 security attacks.

IPv4 option packets that belong to a prefix-based FEC, yet are forwarded into an IPv4/MPLS network without being MPLS-encapsulated, present a security risk against the MPLS infrastructure. Further, LERs that are unable to MPLS encapsulate IPv4 packets with header options cannot operate as an LER in certain MPLS environments. This new requirement will reduce the risk of IPv4 options-based security attacks against LSRs as well as assist LER operation across MPLS networks that minimize the IPv4 routing information (e.g., no BGP transit routes) carried by LSRs.

3. Specification of Requirements

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].

4. Ingress Label Edge Router Requirement

An ingress LER MUST implement the following policy:

- When determining whether to push an MPLS label stack onto an IPv4 packet, the determination is made without considering any IPv4 options that may be carried in the IPv4 packet header. Further, the label values that appear in the label stack are determined without considering any such IPv4 options.

This policy MAY be configurable on an ingress LER, however, it SHOULD be enabled by default. When processing of signaling messages or data packets with more specific forwarding rules is enabled, this policy SHOULD NOT alter the specific processing rules. This applies to, but is not limited to, Resource Reservation Protocol (RSVP) as per [RFC2205], source routing as per [RFC791], as well as other FEC elements defined by future specifications. Further, how an ingress LER processes the IPv4 header options of packets before MPLS encapsulation is out of scope since these are processed before they enter the MPLS domain.

Implementation of the above policy prevents IPv4 packets that belong to a prefix-based FEC from bypassing MPLS encapsulation due to header options. The policy also prevents specific option types such as Router Alert (option value 148) from forcing MPLS imposition of the MPLS Router Alert Label (label value 1) at ingress LERs. Without this policy, the MPLS infrastructure is exposed to security attacks using legitimate IPv4 packets crafted with header options. Further, LERs that are unable to MPLS encapsulate IPv4 packets with header options cannot operate as an LER in certain MPLS environments as described in Section 2.

5. Security Considerations

There are two potential categories of attacks using crafted IPv4 option packets that threaten existing MPLS infrastructures. Both are described below. To mitigate the risk of these specific attacks, the ingress LER policy specified above is required.

5.1. IPv4 Option Packets That Bypass MPLS Encapsulation

Given that a router's exception handling process (i.e., CPU, processor line-card bandwidth, etc.) used for IPv4 header option processing is often shared with IPv4 control and management protocol router resources, a flood of IPv4 packets with header options may adversely affect a router's control and management protocols, thereby, triggering a denial-of-service (DoS) condition. Note, IPv4 packets with header options may be valid transit IPv4 packets with legitimate sources and destinations. Hence, a DoS-like condition may be triggered on downstream transit IPv4 routers that lack protection mechanisms even in the case of legitimate IPv4 option packets.

IPv4 option packets that belong to a prefix-based FEC yet bypass MPLS encapsulation at an ingress LER may be inadvertently IPv4 routed downstream across the MPLS core network (not label switched). This allows an external attacker the opportunity to maliciously craft seemingly legitimate IPv4 packets with specific IPv4 header options in order to intentionally bypass MPLS encapsulation at the MPLS edge (i.e., ingress LER) and trigger a DoS condition on downstream LSRs. Some of the specific types of IPv4 option-based security attacks that may be leveraged against MPLS networks include the following:

- Crafted IPv4 option packets that belong to a prefix-based FEC yet bypass MPLS encapsulation at an ingress LER may allow an attacker to DoS downstream LSRs by saturating their software forwarding paths. By targeting a LSR's exception path, control and management protocols may be adversely affected and, thereby, an LSR's availability. This assumes, of course, that downstream LSRs lack protection mechanisms.

- Crafted IPv4 option packets that belong to a prefix-based FEC yet bypass MPLS encapsulation at an ingress LER may allow for IPv4 Time to Live (TTL) expiry-based DoS attacks against downstream LSRs. MPLS enables IPv4 core hiding whereby transit IPv4 traffic flows see the MPLS network as a single router hop [RFC3443]. However, MPLS core hiding does not apply to packets that bypass MPLS encapsulation and, therefore, IPv4 option packets may be crafted to expire on downstream LSRs which may trigger a DoS condition. Bypassing MPLS core hiding is an additional security consideration since it exposes the network topology.

- Crafted IPv4 option packets that belong to a prefix-based FEC yet bypass MPLS encapsulation at an ingress LER may allow for DoS attacks against downstream LSRs that do not carry the IPv4 routing information required to forward transit IPv4 traffic. Lack of such IPv4 routing information may prevent legitimate IPv4 option packets from transiting the MPLS network and, further, may trigger generation of ICMP destination unreachable messages, which could lead to a DoS condition. This assumes, of course, that downstream LSRs lack protection mechanisms and do not carry the IPv4 routing information required to forward transit traffic.

- Crafted IPv4 option packets that belong to a prefix-based FEC yet bypass MPLS encapsulation at an ingress LER may allow an attacker to bypass LSP Diffserv tunnels [RFC3270] and any associated MPLS Class of Service (CoS) field [RFC5462] marking policies at ingress LERs and, thereby, adversely affect (i.e., DoS) high-priority traffic classes within the MPLS core. Further, this could also lead to theft of high-priority services by unauthorized parties. This assumes, of course, that the [RFC3270] Pipe model is deployed within the MPLS core.

- Crafted RSVP packets that belong to a prefix-based FEC yet bypass MPLS encapsulation at an ingress LER may allow an attacker to build RSVP soft-states [RFC2205] [RFC3209] on downstream LSRs which could lead to theft of service by unauthorized parties or to a DoS condition caused by locking up LSR resources. This assumes, of course, that the MPLS network is enabled to process RSVP packets.

The security attacks outlined above specifically apply to IPv4 option packets that belong to a prefix-based FEC yet bypass ingress LER label stack imposition. Additionally, these attacks only apply to IPv4 option packets forwarded using the global routing table (i.e., IPv4 address family) of a ingress LER. IPv4 option packets associated with a BGP/MPLS IPv4 VPN service are always MPLS

encapsulated by the ingress LER per [RFC4364] given that packet forwarding uses a Virtual Forwarding/Routing (VRF) instance. Therefore, BGP/MPLS IPv4 VPN services are not subject to the threats outlined above [RFC4381]. Further, IPv6 [RFC2460] makes use of extension headers not header options and is therefore outside the scope of this document. A separate security threat that does apply to both BGP/MPLS IPv4 VPNs and the IPv4 address family makes use of the Router Alert Label. This is described directly below.

5.2. Router Alert Label Imposition

[RFC3032] defines a Router Alert Label (label value of 1), which is analogous to the Router Alert IPv4 header option (option value of 148). The MPLS Router Alert Label (when exposed and processed only) indicates to downstream LSRs to examine these MPLS packets more closely. MPLS packets with the MPLS Router Alert Label are also handled as an exception by LSRs and, again, in a less efficient manner. At the time of this writing, the only legitimate use of the Router Alert Label is for LSP ping/trace [RFC4379]. Since there is also no formal standard for Router Alert Label imposition at ingress LERs:

- Crafted IPv4 packets with specific IPv4 header options (e.g., Router Alert) and that belong to a prefix-based FEC may allow an attacker to force MPLS imposition of the Router Alert Label at ingress LERs and, thereby, trigger a DoS condition on downstream LSRs. This assumes, of course, that downstream LSRs lack protection mechanisms.

6. Acknowledgements

The authors would like to thank Adrian Cepleanu, Bruce Davie, Rick Huber, Chris Metz, Pradosh Mohapatra, Ashok Narayanan, Carlos Pignataro, Eric Rosen, Mark Szczesniak, and Yung Yu for their valuable comments and suggestions.

7. References

7.1. Normative References

[RFC791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

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

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

7.2. Informative References

[RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, May 2002.

[RFC3443] Agarwal, P. and B. Akyol, "Time To Live (TTL) Processing in Multi-Protocol Label Switching (MPLS) Networks", RFC 3443, January 2003.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.

[RFC4381] Behringer, M., "Analysis of the Security of BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4381, February 2006.

[RFC4950] Bonica, R., Gan, D., Tappan, D., and C. Pignataro, "ICMP Extensions for Multiprotocol Label Switching", RFC 4950, August 2007.

[IANA] "IP Option Numbers," IANA, February 15, 2007, <www.iana.org>.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, February 2009.

Authors' Addresses

David J. Smith
Cisco Systems
111 Wood Avenue South
Iselin, NJ 08830
EMail: djsmith@cisco.com

John Mullooly
Cisco Systems
111 Wood Avenue South
Iselin, NJ 08830
EMail: jmullool@cisco.com

William Jaeger
AT&T
200 S. Laurel Avenue
Middletown, NJ 07748
EMail: wjaeger@att.com

Tom Scholl
nLayer Communications
209 West Jackson, Suite 700
Chicago, IL 60606
EMail: tscholl@nlayer.net