Network Working Group
Request for Comments: 4182
Updates: 3032
Category: Standards Track

E. Rosen
Cisco Systems, Inc.

September 2005

Removing a Restriction on the use of MPLS Explicit NULL

Status of This Memo

Copyright Notice

Abstract

The label stack encoding for Multi-protocol Label Switching (MPLS) defines a reserved label value known as "IPv4 Explicit NULL" and a reserved label value known as "IPv6 Explicit NULL". Previously, these labels were only legal when they occurred at the bottom of the MPLS label stack. This restriction is now removed, so that these label values may legally occur anywhere in the stack.

This document updates RFC 3032.

Table of Contents

1. Introduction

RFC 3032 defines a reserved label value known as "IPv4 Explicit NULL" and a reserved label value known as "IPv6 Explicit NULL" [RFC3032]. It states that these label values are only legal at the bottom of the MPLS label stack. However, no reason is given for this restriction.

It has turned out that in practice there are some situations in which it is useful to send MPLS packets that have Explicit NULL occur somewhere other than at that bottom of the label stack. While the intended semantics are obvious enough, the fact that such packets are gratuitously declared by RFC 3032 to be illegal has made it difficult to handle these situations in an interoperable manner.

This document updates RFC 3032 by removing the unnecessary restriction, so that the two aforementioned label values are legal anywhere in the label stack.

2. Detail of Change

RFC 3032 states on page 4:

There are several reserved label values:

i. A value of 0 represents the "IPv4 Explicit NULL Label". This label value is only legal at the bottom of the label stack. It indicates that the label stack must be popped, and the forwarding of the packet must then be based on the IPv4 header.

iii. A value of 2 represents the "IPv6 Explicit NULL Label". This label value is only legal at the bottom of the label stack. It indicates that the label stack must be popped, and the forwarding of the packet must then be based on the IPv6 header.

Paragraph i is hereby changed to read:

i. A value of 0 represents the "IPv4 Explicit NULL Label".

An IPv4 Explicit NULL at the top of the label stack means that the stack must be popped.

If the NULL was not the only label on the stack, this will cause the label beneath it to rise to the top of the stack. The disposition of the packet is based on the label that has now risen to the top.

If, on the other hand, the NULL was the only label on the stack, then the stack is now empty. The resulting packet is treated as an IPv4 packet, and its disposition is based on the IPv4 header.

Paragraph iii is hereby changed to read:

iii. A value of 2 represents the "IPv6 Explicit NULL Label".

An IPv6 Explicit NULL at the top of the label stack means that the stack must be popped.

If the NULL was not the only label on the stack, this will cause the label beneath it to rise to the top of the stack. The disposition of the packet is based on the label that has now risen to the top.

If, on the other hand, the NULL was the only label on the stack, then the stack is now empty. The resulting packet is treated as an IPv6 packet, and its disposition is based on the IPv6 header.

3. Reasons for Change

Restricting Explicit NULL to the bottom of the stack has caused some problems in practice.

With this restriction in place, one should not distribute, to a particular label distribution peer, a binding of Explicit NULL to a particular Forwarding Equivalence Class (FEC), unless the following condition (call it "Condition L") holds: all MPLS packets received by that peer with an incoming label corresponding to that FEC contain only a single label stack entry. If Explicit NULL is bound to the FEC, but Condition L doesn't hold, the peer is being requested to create illegal packets. None of the MPLS specifications say what the peer is actually supposed to do in this case. This situation is made more troublesome by the facts that, in practice, Condition L rarely holds, and it is not possible, in general, to determine whether it holds or not.

Further, if one is supporting the Pipe Model of RFC 3270 [RFC3270], there are good reasons to create label stacks in which Explicit NULL is at the top of the label stack, but a non-null label is at the bottom.

RFC 3270 specifies the procedures for MPLS support of Differentiated Services. In particular, it defines a "Pipe Model" in which (quoting from RFC 3270, Section 2.6.2):

"tunneled packets must convey two meaningful pieces of Diff-Serv information:

- the Diff-Serv information which is meaningful to intermediate nodes along the LSP span including the LSP Egress (which we refer to as the 'LSP Diff-Serv Information'). This LSP Diff-Serv Information is not meaningful beyond the LSP Egress: Whether Traffic Conditioning at intermediate nodes on the LSP span affects the LSP Diff-Serv information or not, this updated Diff-Serv information is not considered meaningful beyond the LSP Egress and is ignored.

- the Diff-Serv information which is meaningful beyond the LSP Egress (which we refer to as the 'Tunneled Diff-Serv Information'). This information is to be conveyed by the LSP Ingress to the LSP Egress. This Diff-Serv information is not meaningful to the intermediate nodes on the LSP span."

When the Pipe Model is in use, it is common practice for the LSP Egress to bind Explicit Null to the tunnel's FEC. The intention is that the LSP Diff-Serv information will be carried in the EXP bits of the Explicit Null label stack entry, and the tunneled Diff-Serv information will be carried in whatever is "below" the Explicit Null label stack entry, i.e., in the IP header DS bits or in the EXP bits of the next entry on the MPLS label stack.

Naturally, this practice causes a problem if the Pipe Model LSP is being used to tunnel MPLS packets (i.e., if Condition L does not hold). With strict adherence to RFCs 3031 and 3036, this practice results in an MPLS packet where Explicit NULL is at the top of the label stack, even though it is not the only entry in the label stack. However, RFC 3032 makes this packet illegal.

Some implementations simply transmit the illegal packet. Others try to convert it to a legal packet by stripping off the Explicit NULL before transmitting it. However, that breaks the Pipe Model by discarding the LSP Diff-Serv information. It is conceivable that there may be an implementation that drops the illegal packet entirely; this would also break the Pipe Model, as it would lose not only the LSP Diff-Serv information, but the entire packet.

Of course the LSP egress is not compelled to bind Explicit NULL to the tunnel's FEC; an ordinary label could be used instead. However, using Explicit NULL enables the egress to determine immediately (i.e., without need for lookup in the Label Information Base) that the further forwarding of the packet is to be determined by whatever is below the label. Avoiding this lookup can have favorable implications on forwarding performance.

Removing the restriction that Explicit Null only occur at the bottom of the stack is the simplest way to facilitate the proper operation of the Pipe Model.

4. Deployment Considerations

Implementations that adhere to this specification will interoperate correctly, and will correctly support the Pipe Model.

Implementations that do not adhere to this specification may not interoperate. In particular, if a router advertises a binding of Explicit NULL, and if that router has an upstream LDP peer that will not transmit a packet that has multiple label stack entries with Explicit Null at top of the stack, then it will not be possible to use Explicit NULL to support the Pipe Model until the upstream LDP peer is brought into compliance with this specification.

It is possible that there may be a router implementation, preceding this specification, which will discard any received packet with multiple label stack entries and a top label value of Explicit Null. It is advisable to configure any such routers so that they do not advertise any bindings to Explicit Null.

5. Security Considerations

This document updates RFC 3032 by allowing Explicit NULL to occur at any position in the label stack. This modification does not impose any new security considerations beyond those discussed in RFC 3032.

6. Acknowledgments

Thanks to Rahul Aggarwal, Francois LeFaucheur, Yakov Rekhter, and Dan Tappan for their helpful comments.

7. Normative References

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

8. Informative References

[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.

Author's Address

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

EMail: erosen@cisco.com

Full Copyright Statement

Intellectual Property

Acknowledgement

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