

Network Working Group  
Internet-Draft  
Updates: [3032](#) (if approved)  
Intended status: Standards Track  
Expires: January 1, 2016

Z. Zhang  
Juniper Networks, Inc.  
June 30, 2015

MPLS ICMP for BIER Payload  
draft-zzhang-mpls-icmp-bier-01.txt

## Abstract

This document specifies an optional extension to generate ICMP messages for BIER packets transported over LSPs.

## Requirements Language

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](#).

## Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on January 1, 2016.

## Copyright Notice

Copyright (c) 2015 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 (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

Internet-Draft

mpls-icmp-bier

June 2015

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">2.</a>	Specifications . . . . .	<a href="#">4</a>
<a href="#">3.</a>	IANA Considerations . . . . .	<a href="#">5</a>
<a href="#">4.</a>	Security Considerations . . . . .	<a href="#">5</a>
<a href="#">5.</a>	Acknowledgements . . . . .	<a href="#">5</a>
<a href="#">6.</a>	References . . . . .	<a href="#">5</a>
<a href="#">6.1.</a>	Normative References . . . . .	<a href="#">5</a>
<a href="#">6.2.</a>	Informative References . . . . .	<a href="#">6</a>
	Author's Address . . . . .	<a href="#">6</a>

## [1.](#) Introduction

[RFC3032](#) specifies that an LSR may generate ICMP messages if the payload is IPv4 or IPv6. Normally the ICMP messages are label switched using the original label stack, and the egress LSR will then forward the the messages to the source of the packet.

BIER [wijnands-bier-architecture] is a new multicast forwarding architecture and [kumarzheng-bier-ping] specifies ping/traceroute procedures for BIER. BIER traceroute uses the same TTL-expiration principle of IPv4/IPv6 unicast traceroute. A BFR (a router that supports BIER) may get a traceroute probe message whose TTL just expired at this hop and may send back a response, allowing a BFIR (ingress BFR) to gather the paths that a BIER packet traverses.

It's possible that a BIER packet is transported over an LSP between two BFRs. If Uniform Model for TTL is used for the LSP, when the ingress LER for the LSP put the BIER packet with a preceeding BIER label into the tunnel, the TTL from the BIER label is copied to the outgoing label for the LSP. As a result, the TTL expiration may happen on an LSR on the LSP, who will silently drop the packet. This TTL expiration could easily happen with traceroute, because the BFR that initiates the tracing will increase the TTL to be used one by one to discover the path. The silent drop by an LSR is therefore undesired.

If the LSR can generate an ICMP message for the TTL expiration, label switch it to the LER that advertised the inner most label just like in the IPv4/IPv6 case, that LER, which is is a BFR, will be able to process the TTL expiration for BIER traceroute purpose.

Note that in IPv4/IPv6 case, the generated ICMP message is addressed to the packet's source address and the message will be transparently routed back to the source of the packet by the LER that advertised the inner most label in the stack. For BIER, the ICMP message is to be processed by the LER that advertised the inner most label. Therefore, the destination address of the ICMP message is set to local host address 127.0.0.1 or ::1, depending on if the MPLS infrastructure is based on IPv4 or IPv6.

In the following example diagram, there is an ingress BFR (BFIR), an egress BFR (BFER), and two transit BFRs (BFR1/BFR2) separated by two non-BFRs. When BFIR sends a BIER packet, BFR1 will put the BIER packets into a tunnel between BFR1 and BFR2. If Uniform model is used on BFR1, the tunneled packet could have TTL expiration on non-BFR1. When that happens, non-BFR1 will generate an ICMP message addressed to local host address 127.0.0.1 or ::1, and label switch to BFR2. BFR2 will then process the ICMP message and may send appropriate response to the BFIR.

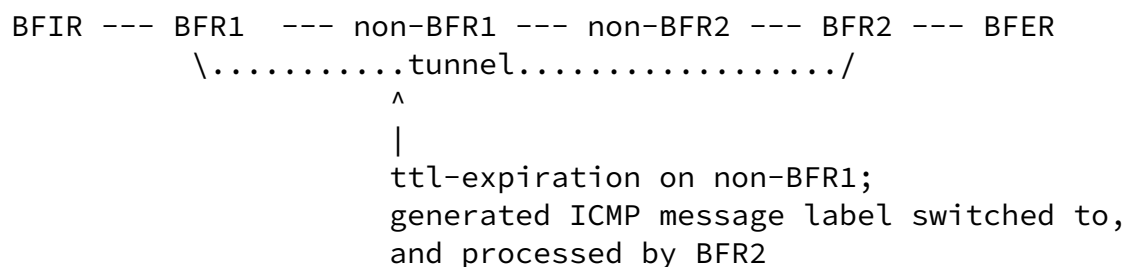


Figure 1

In theory, it is possible that the outer LSP for which the LSR is having TTL expiration is the base LSP for a stacked LSP and the latter uses a differently versioned IP infrastructure, so the version of the ICMP message may be chosen incorrectly. In reality, this is unlikely to happen because the label stack is to transport BIER packets between two BFRs that are in the same IGP domain. Even if it does happen, it is likely that the LER is able to process both ICMP

and ICMPv6 messages, and even if the LER is not able to process the incoming ICMP/ICMPv6 message it can discard the message silently - not much different from the LSR not generating the message the first place.

It is expected that the BIER encapsulation header will start with a 4-bit nibble with a unique value that suggests that it is a BIER packet.

This specification focuses on ICMP message generated for TTL expiration. Other types of ICMP messages are out of the scope at this time.

## [2.](#) Spefications

When an LSR experiences a TTL expiration for a labeled packet and the first 4-bit nibble is X [TBD], the LSR MAY generate an ICMP message if the MPLS infrastructure is IPv4 based, or an ICMPv6 message if the MPLS infrastructure is IPv6 based. The destination of the message is 127.0.0.1 in case of IPv4 or ::1 in case of IPv6. The source of the message is a routable address of the LSR.

The LSR, if it supports BIER, MAY further check the BIER payload type. If the "proto" field of the BIER header is "BIER OAM", the LSR SHOULD generate an ICMP or ICMPv6 message.

To differentiate from ICMP/ICMPv6 messages generated for IPv4/IPv6 payloads, the following message types are defined:

ICMP messages:

Type TBD1: TTL expired in tunnel

ICMPV6 messages:

Type TBD2: TTL expired in tunnel

Alternatively, same message types but different code could be used to differentiate from IPv4/IPv6 payload triggered messages. [This will be decided based on WG input]

Except for the differences mentioned above in message type/code, destination address, and ICMP/ICMPv6 choice based on MPLS infrastructure type, the ICMP/ICMPv6 messages are constructed following existing procedures for IPv4/IPv6 payload. The generated

messages are label switched using the invoking packet's original label stack minus the inner most label. The original inner most label for a BIER packet is a BIER label indicating that the payload type is BIER, so it must not be used for switching the generated ICMP/ICMPv6 messages, which are not BIER packets. For a true BIER packet, the next inner most label is for the LSP terminating at the BFR that advertised the inner most label, so the original label stack minus the inner most label should get the generated ICMP/ICMPv6 message to the right place for further processing. For the example in Figure 1, when non-BRF1 experiences a TTL-expriation for a tunneled BIER packet, the label stack could be <tunnel label, BIER label>, where the tunnel label is advertised by non-BFR2 and the BIER label is advertised by BFR2. With the BIER label removed, the resulting label stack <tunnel label> will get the generated ICMP message to BFR2 for further processing.

The LER that originated the inner most label for the generated messages will receive the ICMP/ICMPv6 messages, which is addressed to the local host, and process the messages locally.

How the messages are processed is outside of the scope of this document, but in general the new ICMP message type/code causes the LER to examine the original packet header that is enclosed in the ICMP/ICMPv6 message and act accordingly, e.g., forward to BIER OAM module for further processing. Because a non-BIER packet may be mistaken as a BIER packet (if the first nibble happens to be match), the receiving BFR MUST check the label stack included in the ICMP/ICMPv6 message to make sure that the inner most label is a BIER label that it advertised.

It's possible that an IPv4 LER cannot process an incoming ICMPv6 message, or vice versa. It is also possible that an LER cannot recognize the new message type/code. In these cases, it simply discard the message.

### [3.](#) IANA Considerations

This document requests IANA to assigna a new message type in ICMP and ICMPv6 Parameters registries respectively.

### [4.](#) Security Considerations

This document does not introduce new security risks, compared to generating ICMP messages for labeled IP packets.

## [5.](#) Acknowledgements

The author thanks Eric Rosen, Ronald Bonica for their review, comments, and suggestions.

## [6.](#) References

### [6.1.](#) Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", [RFC 3032](#), January 2001.
- [RFC4443] Conta, A., Deering, S., and M. Gupta, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", [RFC 4443](#), March 2006.
- [RFC4884] Bonica, R., Gan, D., Tappan, D., and C. Pignataro, "Extended ICMP to Support Multi-Part Messages", [RFC 4884](#), April 2007.

Zhang

Expires January 1, 2016

[Page 5]

---

Internet-Draft

mpls-icmp-bier

June 2015

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

### [6.2.](#) Informative References

- [I-D.ietf-bier-architecture]  
Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", [draft-ietf-bier-architecture-00](#) (work in progress), April 2015.
- [I-D.kumarzheng-bier-ping]  
Kumar, N., Pignataro, C., Akiya, N., Zheng, L., Chen, M., and G. Mirsky, "BIER Ping and Trace", [draft-kumarzheng-](#)

[bier-ping-00](#) (work in progress), March 2015.

Author's Address

Zhaohui Zhang  
Juniper Networks, Inc.  
10 Technology Park Drive  
Westford, MA 01886

EMail: [zzhang@juniper.net](mailto:zzhang@juniper.net)