

Bidirectional Forwarding Detection  
Internet-Draft  
Intended status: Informational  
Expires: January 12, 2012

J. Haas, Ed.  
Juniper Networks  
M. Xiao, Ed.  
ZTE Corporation  
July 11, 2011

Application of the BFD Echo function for Path MTU Verification or  
Detection  
draft-haas-xiao-bfd-echo-path-mtu-01

## Abstract

This document specifies an extended application of the BFD Echo function for path MTU verification or detection, while preserving its original purpose for detecting forwarding failures. This document defines a process to vary the length of some BFD Echo packets periodically to accomplish this Path MTU verification or detection.

## 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 12, 2012.

## Copyright Notice

Copyright (c) 2011 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 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="#">3</a> |
| <a href="#">2.</a>     | Conventions . . . . .  | <a href="#">3</a> |
| <a href="#">3.</a>     | Abbreviations . . . . .  | <a href="#">3</a> |
| <a href="#">4.</a>     | Deployment Scenario . . . . .  | <a href="#">4</a> |
| <a href="#">5.</a>     | Format of the BFD Echo Packets for Path MTU Verification<br>or Detection . . . . . | <a href="#">4</a> |
| <a href="#">6.</a>     | Extension to BFD Echo Operation . . . . .  | <a href="#">4</a> |
| <a href="#">6.1.</a>   | Verification of Path MTU . . . . .   | <a href="#">5</a> |
| <a href="#">6.1.1.</a> | BFD Echo Packet Transmission . . . . .   | <a href="#">5</a> |
| <a href="#">6.1.2.</a> | BFD Echo Packet Reception . . . . .  | <a href="#">5</a> |
| <a href="#">6.2.</a>   | Detection of Path MTU . . . . .  | <a href="#">6</a> |
| <a href="#">6.2.1.</a> | BFD Echo Packet Transmission . . . . .   | <a href="#">6</a> |
| <a href="#">6.2.2.</a> | BFD Echo Packet Reception . . . . .  | <a href="#">7</a> |
| <a href="#">6.2.3.</a> | Consequent Actions . . . . .   | <a href="#">7</a> |
| <a href="#">7.</a>     | Security Considerations . . . . .  | <a href="#">8</a> |
| <a href="#">8.</a>     | IANA Considerations . . . . .  | <a href="#">8</a> |
| <a href="#">9.</a>     | Acknowledgements . . . . .   | <a href="#">8</a> |
| <a href="#">10.</a>    | References . . . . .   | <a href="#">8</a> |
| <a href="#">10.1.</a>  | Normative References . . . . .   | <a href="#">8</a> |
| <a href="#">10.2.</a>  | Informative References . . . . .   | <a href="#">8</a> |

## 1. Introduction

BFD ([\[RFC5880\]](#)) defines an Echo function as an adjunct to the two operating modes of BFD. When the Echo function is active, a stream of BFD Echo packets is transmitted in such a way as to have the other system loop them back through its forwarding path. If a number of packets of the echoed data stream are not received, to be clearer, if the number of unreceived consecutive Echo packets is more than the value contained in "Detection Multiplier" of the last received BFD Control packet, the BFD session is declared to be down.

As also indicated in [\[RFC5880\]](#), "the Echo function has the advantage of truly testing only the forwarding path on the remote system. This may reduce round-trip jitter and thus allow more aggressive Detection Times, as well as potentially detecting some classes of failure that might not otherwise be detected".

This document specifies an extended application of the BFD Echo function for path MTU verification or detection, while preserving its original purpose for detecting forwarding failures. This document defines a process to vary the length of some BFD Echo packets periodically to accomplish this Path MTU verification or detection.

## 2. Conventions

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 [\[RFC2119\]](#).

## 3. Abbreviations

BFD: Bidirectional Forwarding Detection

ICMPv4: Internet Control Message Protocol version 4

ICMPv6: Internet Control Message Protocol version 6

LDP: Label Distribution Protocol

LSP: Label Switched Path

MTU: maximum Transmission Unit

MPLS: Multiprotocol Label Switching

PE: Provider Edge

RSVP: Resource Reservation Protocol

#### [4.](#) Deployment Scenario

According to [[RFC5880](#)] and [[RFC5884](#)], the BFD Echo function may be deployed for MPLS LSPs. The extended application of the BFD Echo function for Path MTU verification or detection, which is described in this draft, is also expected to be deployed for MPLS LSPs.

The extended application of BFD Echo function can be used either if the MPLS LSP was created statically by manual configuration, or if the MPLS LSP was setup dynamically using a control protocol (e.g. LDP or RSVP).

Note that this draft assumes path MTUs of the forward and return are symmetric. This is because by using the BFD Echo function only the minimum of the path MTUs for the two directions can be verified or detected.

#### [5.](#) Format of the BFD Echo Packets for Path MTU Verification or Detection

As indicated in [[RFC5880](#)], "BFD Echo packets are sent in an encapsulation appropriate to the environment. The payload of a BFD Echo packet is a local matter, since only the sending system ever processes the content. The only requirement is that sufficient information is included to demultiplex the received packet to the correct BFD session after it is looped back to the sender. Some form of authentication SHOULD be included, since Echo packets may be spoofed".

This document requires the following information to be encoded in the BFD Echo packet: The 4-octet BFD My Discriminator field. The 4-octet BFD Your Discriminator field. The BFD optional Authentication Section if the BFD session has set the Authentication Present flag. A variable length padding field, the value of which is not defined by this specification.

The discriminator and authentication fields provide sufficient information for the session to be de-multiplexed upon receipt by the BFD session. The padding field provides the ability to test the forwarding path to carry packets of a given size. The length of the padding field should account for the overhead of the Echo packet's encapsulation type.

## [6.](#) Extension to BFD Echo Operation

As specified in [[RFC5880](#)], one end system of the established BFD session may transmit BFD Echo packets if the last BFD Control packet received from the remote system contains a nonzero value in "Required

Min Echo RX Interval" and the bfd.SessionState is Up. The interval between transmitted BFD Echo packets is set to the higher one of received "Required Min Echo RX Interval" and the minimum sending period supported by the transmitting system, except that a 25% jitter may be applied to the rate of transmission, such that the actual interval may be between 75% and 100% of the advertised value.

As also indicated in [[RFC5880](#)], the remote system of the established BFD session will loop all received BFD Echo packets back to the local system, and if there are consecutive more than N (N equals the value contained in "Detection Multiplier" of the last received BFD Control packet) Echo packets not received at the local system, it's judged that a forwarding failure is detected and the BFD session is declared to be down.

Some extensions to the BFD Echo operation are needed for the extended application defined in this document. Note that these extended operations will not disrupt the existing application of BFD Echo function to detect forwarding failure of the bidirectional transport path.

### [6.1.](#) Verification of Path MTU

After a MPLS LSP is setup, there exists a target path MTU for that LSP. The BFD Echo function can be used to verify the validity of the target path MTU for that LSP.

#### [6.1.1.](#) BFD Echo Packet Transmission

When transmitting BFD Echo packets, the local system should transmit two kinds of BFD Echo packets alternatively: Unpadded echo packets and padded echo packets. Provided the BFD Detection Multiplier is large enough, the unpadded echo packets are sufficient to keep the BFD Session in the Up state even if the padded packets are dropped due to a Path MTU size failure. Similarly, transmission of the padded BFD Echo packets is used to test desired Path MTU.

#### [6.1.2.](#) BFD Echo Packet Reception

When receiving BFD Echo packets to achieve forwarding failure detection and path MTU verification, the local system should first demultiplex the received packet to the correct BFD session using the embedded BFD discriminator fields. If Authentication is present, the Authentication procedure should also be applied to the received BFD Echo packets. The procedure for detecting a forwarding failure in [[RFC5880](#)] is carried out normally. Additionally, if more than Detection Multiplier consecutive padded Echo packets (i.e. every alternate packet) is not received, the Path MTU is considered to be

down. This MAY trigger the detection of the new effective Path MTU.

### [6.2.](#) Detection of Path MTU

For the purpose of determining effective Path MTU, a maximum packet length and a minimum packet length of BFD Echo packet should be configured. For example, the target path MTU could be used as the maximum packet length. The minimum packet length is the minimum required path MTU for the applications carried on the LSP. If this minimum value is unknown, the minimum packet length may be configured to the minimum packet length required for the underlying encapsulation type of the BFD Echo Packet.

#### [6.2.1.](#) BFD Echo Packet Transmission

When transmitting BFD Echo packets to detect both forwarding failure and path MTU, the local system should consecutively transmit BFD Echo packets which are grouped by value N (N equals the value contained in "Detection Multiplier" of the last received BFD Control packet). For the tolerance of possible temporary loss of BFD Echo packets, N MUST be no less than 3. In every group of N Echo packets, the 2nd and the (N-1)th Echo packets should be with a padded Echo packet where the packet length is of a size used to execute a probe operation on the forwarding path.

There are two options that may be used for determining the method of selecting the size of the Path MTU probe packets during the Path MTU detection procedure:

1. The probe packets are set to a length initialized to the minimum packet length required to be supported by the forwarding path. The value is then increased by a step interval that is user configured until the length of the probe packets reaches the maximum packet length.
2. The probe packets are set to a length which provides a binary search of the minimum and the maximum desired packet length. Initially, the minimum packet length is probed. If the forwarding path supports the minimum desired packet length then the maximum packet length is probed. If the probe of the maximum packet length fails, the probe packet size is set to the halfway point between the minimum and maximum packet length and so on per the standard binary search algorithm. In this manner, the effective Path MTU may be determined.

#### [6.2.2.](#) BFD Echo Packet Reception

The procedure for verifying forwarding detection failures should be followed as per the prior section on verifying path MTU. During Path MTU probe operations, the reception of the different sized padded Echo packets is used as inputs for the probing procedure per the transmission procedures above. The reception of a single padded packet of the probe size is considered sufficient for validation of

the probed MTU for that size probe packet. If N consecutive Detection Multiplier probe packets are not detected, the probe for that size packet is considered a failure and the probing procedure reacts accordingly.

### [6.2.3.](#) Consequent Actions

After the process of detecting path MTU finished, there are two possible results: One is that the new effective path MTU is detected, the other is that the new effective path MTU can't be detected because it's below the minimum required path MTU for the application carried on the LSP. Different consequent actions would be taken due to the results.

If the new effective path MTU is detected, it would be reported to the operator. As specified in [section 3 of \[RFC3032\]](#), the detected path MTU of MPLS LSP MAY be used to dynamically determine the maximum size for fragmentation. It should also be noted that for the MPLS LSPs the potential fragmentation would take place on the inner IP datagram and after that the MPLS label stack entries are appended. Also note that fragmentation and reassembly in network equipment generally requires significantly greater resources than sending a packet as a single unit, so fragmentation and reassembly should be avoided whenever possible. For this reason, when the local system (e.g. an ingress PE) receives a packet which is too big to be encapsulated and transmitted as a single unit over the transport path - i.e. the length of encapsulated packet exceeds the detected path MTU - another approach is to discard the received packet and use techniques (e.g. ICMPv4/ICMPv6) to signal the sources whose packets will be encapsulated in the network to send smaller packets.

If the minimum required path MTU for application carried on the LSP is pre-provisioned as the min packet length and the new effective path MTU can't be detected, the consequent action would be to tear the BFD session down just as forwarding failure is detected by the existing application of BFD Echo function. This may trigger protection switching of the LSP.

## [7.](#) Security Considerations



To be added in a later version of this document.

## [8.](#) IANA Considerations

This document introduces no considerations for IANA.

## [9.](#) Acknowledgements

The editors would like to thank Lei Zhang and Xuehui Dai of ZTE Corporation for their valuable input.

## [10.](#) References

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

- [RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", [RFC 5880](#), June 2010.
- [RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", [RFC 5884](#), June 2010.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.

### [10.2.](#) Informative 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.

## Authors' Addresses

Jeffrey Haas (editor)  
Juniper Networks

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

Min Xiao (editor)  
ZTE Corporation

EMail: [xiao.min2@zte.com.cn](mailto:xiao.min2@zte.com.cn)