Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: November 6, 2014

B. Snyder
iDirect Technologies
N. Akiya
Cisco Systems

May 5, 2014

# BFD Proxy for Connections over Monitored Links
## draft-snyder-bfd-proxy-connections-monitored-links-00

## Abstract

This document describes a Bidirectional Forwarding Detection (BFD) proxy mechanism to allow intermediate networking equipment (ex: Satellite HUB/Modem) to intercept BFD packets and to generate BFD packets to relay the health of connection monitored links.

Note that this is an informational document that does not propose any changes to the BFD protocol.

## 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 RFC 2119 [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 November 6, 2014.

## Copyright Notice

Table of Contents



# 1. Introduction

## 1.1. Terminology

The following acronyms/terminologies are used in this document:

- BFD - Bidirectional Forwarding Detection
- DLEP - Dynamic Link Exchange Protocol
- L2 - Layer 2
- L3 - Layer 3
- Outroute - The broadcast link from hub to modem(s) in a satellite network.

- Downstream - Synonymous to Outroute.
- OTA - Over the Air
- Inroute - The unicast uplink that a modem transmits to the hub side on in a satellite network.
- Upstream - Synonymous to Inroute.

## 1.2. Background

Bidirectional Forwarding Detection (BFD) is an application agnostic and link type independent keep alive protocol which has widely been implemented and deployed. The BFD protocol can be configured with a fast interval to provide rapid failure detection or configured with a slower interval to provide slower failure detection. The faster the interval, the more BFD packets are transmitted and received, consuming more system and network resources.

Some links have connection monitoring functionality of its own, and some of these connection monitored links have constraints (ex: limited or expensive bandwidth). Applications over such links often still desire rapid failure detection through exchanging keep-alive packets (ex: BFD). However, the consequence of such can significantly degenerate the value of the links. For example, running BFD over a link with limited bandwidth can result in a significant portion of the bandwidth being consumed by BFD packets.

One example of such scenario is:

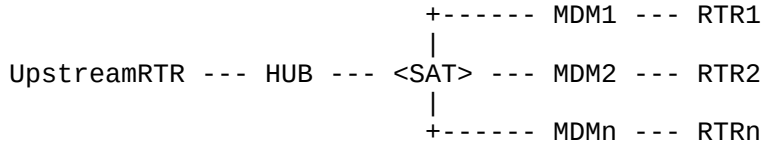


Figure 1: Star Satellite Network

The HUB components consist of a protocol stack which processes and inspects all outbound packets in order to optimize traffic for a high delay low bandwidth environments. (Ex: TCP proxy, compression, encryption). This stack also contains a L2 switch to demultiplex outroute traffic towards the proper modem via a MAC learning switch. In this component is also station keeping algorithms and QOS schedulers.

The MDM components have the same protocol stack (without the demultiplexing required) to optimize the traffic flow for the TDMA

inroutes. The interfaces of a modem are 1 RF interface and 1 to 8 ethernet interfaces.

When routers connected to HUB and MDMs run BFD to monitor the L3 reachability through the Satellite network, expensive Satellite bandwidth gets consumed with large number of BFD packets traversing over it.

Dynamic Link Exchange Protocol (DLEP), [I-D.ietf-manet-dlep], tackles this problem by introducing a protocol that can communicate the state of monitored links to routing devices. DLEP also maintains and communicates an extensive set of information (ex: link quality). A wide range of DLEP responsibilities result in a large effort for vendors to develop this protocol. DLEP, in addition, will require further effort to get integrated into various applications (ex: IGP) for the information to be beneficial.

BFD, on the other hand, has widely been implemented and deployed. If applications, already capable of speaking BFD, only require keeping track of a connection state over monitored links, and not any other information provided by DLEP, then the BFD proxy, described in this document, can be implemented on intermediate networking equipment to allow:

- Connected network equipment (i.e. routers) to continue using BFD for continuity check.

- BFD packets to consume minimal bandwidth on monitored links.

## 2. Overview

This informational document describes a BFD proxy mechanism that allows for connection monitoring intermediate networking equipment (ex: Satellite HUB/Modem) to use BFD packets in order to communicate the state of monitored links whilst significantly reducing the network bandwidth consumed by BFD packets.

The BFD proxy is a link state aware module that resides on the intermediate networking equipment, and intercepts all BFD packets coming in from the connected network equipment.

The first task of the BFD proxy is to transmit BFD control packets to the connected network equipment in order to communicate the state of monitored links, based on its knowledge of link state. The BFD proxy can inject BFD STATE change events towards the connected network equipment. When the device under monitoring is present, the BFD proxy can inject BFD packets with BFD_UP state. When the device

under monitoring has left the network, the BFD proxy can inject BFD packets with BFD_DOWN state.

The second task, to reduce network bandwidth is handled both at the BFD level (by proxing) and at the L3 level. By proxing the BFD control packets one can keep all the BFD overhead off the monitored (and often expensive) bandwidth links. The use of BFD also allows for network designers to configure L3 keep-alive/HELLO timers to be increased thereby reducing OTA bandwidth usage of un-proxied data flows. With BFD monitoring and alerting, L3 convergence is bound by a combination of link state awareness and IGP Hello time (in either direction). The monitored link's state (ex: satellite modem) can be immediately propagated when transitioning between in and out of network. Additionally, configurations and protocols will be discussed that have been determined to be optimal to this use case.

This document will also suggest multiple integration improvements that all interested parties (routing vendors and modem vendors) could implement to further optimize convergence time and bandwidth usage. The network configuration is that of a star design, where thousands of CE routers each behind a satellite remote will attach to one hub upstream router via desired L3 protocols. Whilst, many networks do utilize mobility and roaming, they are always aware of whom they are connecting too (either one or more possible HUBs, but only one at a time). As the goal is simply to assist the routers in understanding radio link state to optimize routing convergence, BFD is the optimal way of meeting this need.

## 3. BFD Proxy Placement

The BFD proxy module MUST be placed on a system such that it meets following two criteria:

1. The BFD proxy module can access the state of monitored links and neighbors reachable through it.

2. The BFD proxy module can access all single-hop BFD control packets coming in from the connected network equipment.

## 4. BFD Proxy Procedures

### 4.1. BFD Control Packet Interception

The BFD proxy module MUST intercept all single-hop BFD control packets (referred to as BFD packets from hereon) coming in from the connected network equipment. Criteria to identify single-hop BFD control packets are:

1. IP/UDP Packet

2. IP TTL 255 ([RFC5881] and for [RFC5082])

3. UDP destination port 3784 ([RFC5881])

### 4.2. OAM Object

The BFD proxy module SHOULD maintain an OAM object per neighbor reachable through monitored links. This OAM object is to have the state of the neighbor (i.e. available or not available), stores local BFD discriminator value and caches the latest BFD packet intercepted. When the BFD proxy module intercepts a BFD packet, destination MAC address is used to locate the OAM object. If corresponding OAM object is not found, then perform local checks to see if one should get created. If the check passes, create the OAM object. Otherwise do not create one.

### 4.3. BFD Proxying

Upon intercepting a BFD packet and locating a corresponding OAM object, the BFD proxy module is to follow procedures described in this sub-section.

1. If there is no OAM object, no further action is taken.

2. If the state of the neighbor in the OAM object is "not-available", then no further action is taken.

3. If the State field of intercepted BFD control packet is:

   * ADMIN_DOWN: Forward the intercepted packet OTA to alert the real destination.

   * DOWN: Create a BFD packet and copy the contents from intercepted packet, with the following modifications:

     + Swap source and destination MAC addresses.

     + Swap source and destination IP addresses.

     + Set "my discriminator" field.

     + Clear "your discriminator" field.

     Send constructed BFD packet to the connected network equipment.

* INIT: If "your discriminator" does not match expected value, then no further action is taken. Otherwise, create a BFD packet and copy the contents from the intercepted packet, with the following modifications:

  + Swap source and destination MAC addresses.

  + Swap source and destination IP addresses.

  + Swap "my discriminator" and "your discriminator" fields.

  + Set "State" field to UP.

  Send constructed BFD packet to the connected network equipment.

* UP: If "your discriminator" does not match the expected value, then no further action is taken. Otherwise. create a BFD packet and copy the contents from the intercepted packet, with following modifications:

  + Swap source and destination MAC addresses.

  + Swap source and destination IP addresses.

  + Swap "my discriminator" and "your discriminator" fields.

  Send constructed BFD packet to the connected network equipment.

In addition, following procedures MAY be applied:

o When a BFD control packet is sent to the connected network equipment, the UDP checksum is set to 0 to avoid the recalculation.

o When the state of the neighbor in the OAM object changes from "available" to "not-available", then the BFD proxy module SHOULD send unsolicited BFD control packet with state field as DOWN to the connected network equipment. If this is not done, then absence of a "reply" BFD control packet from the BFD proxy will cause the sending router to timeout the connection after 3 drops (or whatever the multiplier is set too).

o Once the BFD proxy is intercepting BFD control packets and is in UP state, Poll sequence MAY be initiated to increase values in Minimum TX Interval and Minimum RX Interval fields to reduce the

number of BFD control packets on the link connecting the network equipment and the intermediate network equipment.

- o Since on a Satellite Star Network configuration the outroute and inroute have different bandwidth considerations, there are unique integration concerns which are described below

## 4.4. Outroute Considerations

In a star satellite network, the outroute is a broadcast channel which all remotes receive. While there need not be any restrictions on L3 routing protocols, it does naturally follow that an IGP is a good choice. Specifically, one which allows for asynchronous timers.

Terrestrial convergence timing with BFD (sub second) is in the most common error cases (rainfade, mobility switching) not a realistic goal as the RF algorithms that determine out of network will take on the order of seconds (15 in this specific case). Therefore should a modem leave the network for any reason, the minimum convergence time at the hub side is 15 seconds plus BFD timing to recognize the link loss. Hence, the goal being to minimize bandwidth overhead to make this as short as possible above layer 1 timing. A further consideration is convergence timing when a modem comes back into network. If the L3 timers are made too high, then it can take too long to recognize a positive network state. The outroute being a broadcast medium, can work well within these parameters if for instance the outroute L3 hello timing was every 5 seconds. That’s only 1 multicast hello packet to cover the entire network and will bound the convergence time to within 5 seconds.

## 4.5. Inroute Considerations

On the inroute, network bandwidth is much harder to come by, because the aggregate throughput of all inroutes is shared amongst all modems (potentially numbering in the tens of thousands), and is very expensive. Also, it is unicast to the hub side only. Therefore any decisions made here on timing and data transmissions must scale to the tens of thousands in design principles. This fact is the catalyst for preferring asynchronous timers. Ideally, one can rely on the hello packet of a multicast outroute to kick off convergence, and the hello timing of the inroute can be tuned down as much as possible, to optimize inroute usage. This is possible with EIGRP and IS-IS protocols. Unfortunately, BGP and OSPF require synchronized timers, which means it is impossible to weigh equally the convergence timing while protecting inroute bandwidth.

Additionally, further integration simplicity can optionally be achieved if desired. Notice the timing of 15 seconds to recognize

modem link state is also 3 (a common multiplier setting) times the 5 seconds (common hello message timing). Therefore, it is possible, if one is only interested in monitoring link state, to not utilize BFD on all the remote LANs, as 15 seconds is enough time for the L3 messaging to alert the router to a network issue and just about the same time that the hub side will notice. This is useful to simplify operational complexity and management of the thousands or tens of thousands of installed networks. If one would like BFD to monitor modem LAN state as well, then it would be required regardless.

## 5. Possible Integration Improvements

The following improvements could help with overhead and convergence timing in all monitored network environments. They can require changes on routing or modem equipment to further optimize these types of networks:

- BFD timer - Allowing for connected network equipment to configure a high BFD interval value. One of BFD’s missions is to support sub second failure notification. This document puts forth a useful situation in which BFD is a great help, but does not require such strict timing. In fact, it would scale better with much looser restrictions on timer configuration.

- BFD demand mode implementation - If vendors had implemented demand mode, it would be possible for the BFD proxy to send D bit to the connected network to significantly minimize BFD packets traversing over local link connected to the network equipment, without tweaking Minimum TX Interval and Minimum RX Interval values. This would reduce processing of BFD packets by the BFD proxy module even further.

- BFD protocol - Adding into the core protocol the notion of a proxier could assist with support of authentication in this use case, if desired.

## 6. Security Considerations

The proxying by the BFD proxy module will require additional considerations (i.e. knowing authentication types/keys of each neighbor) to handle BFD packets with BFD authentication data (described in Section 6.7 of [RFC5880]. This document only describes procedures to handle BFD packets without BFD authentication data. However, because the mechanism is only applicable to single-hop BFD ([RFC5881]) and GTSM (i.e. check for TTL=255) already provides fairly strong security, lack of BFD authentication support is not considered threatening.

## 7. IANA Considerations

This document does not define any code points.

## 8. Acknowledgements

Authors would like to thank Adrian Farrel for providing a suggestion to generalize the solution to all monitored links.

## 9. References

### 9.1. Normative References

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

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.

[RFC5881] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6 (Single Hop)", RFC 5881, June 2010.

### 9.2. Informative References

[I-D.ietf-manet-dlep] Ratliff, S., Cisco, C., Harrison, G., Jury, S., and D. Satterwhite, "Dynamic Link Exchange Protocol (DLEP)", draft-ietf-manet-dlep-05 (work in progress), February 2014.

[RFC5082] Gill, V., Heasley, J., Meyer, D., Savola, P., and C. Pignataro, "The Generalized TTL Security Mechanism (GTSM)", RFC 5082, October 2007.

## Authors' Addresses

Brian Snyder
iDirect Technologies

Email: bsnyder@idirect.net

Nobo Akiya
Cisco Systems

Email: nobo@cisco.com