Bidirectional Forwarding Detection (BFD) implementation and support in OpenBSD

Peter Hessler
phessler@openbsd.org

OpenBSD

11 March, 2017

first, some background

- routing is the delivery of packets
- first step: lookup the destination
- we have a route, is it usable?
- ... check the gateway for the RTF_UP state

before bfd

- normally, you monitor the link state
- ...not always reliable
- sometimes there are active devices between you and your neighbor
- ...switches
- ...long reach connect

what is bfd?

- bgp timers are generally 90 seconds
- how much traffic is that when you are sending 10Gbps?
- .... 100Gbps?
- .... your ceo is talking to you over voip

what is bfd?

- bgp timers are generally 90 seconds
- how much traffic is that when you are sending 10Gbps?
- .... 100Gbps?
- .... your ceo is talking to you over voip
- fastest possible bgp holdtime is 3 seconds!

what is bfd?

- bgp timers are generally 90 seconds
- how much traffic is that when you are sending 10Gbps?
- .... 100Gbps?
- .... your ceo is talking to you over voip
- fastest possible bgp holdtime is 3 seconds!
- ospf? ldp? other protocols?

what is bfd?

- bidirectional forwarding detection (RFC 5880)
  - detecting faults between two forwarding devices
  - kinda like gre-keepalives
  - protocol independent
- bfd for ipv4 and ipv6 (single hop) (RFC 5881)
  - encapsulates bfd payload in a normal udp packet

what is bfd?

- found on big routers
- ...commonly used with bgp
- ...or mpls
- specs use microseconds!
- (μs not ms)
- ...implementation detail, we won't support timers faster than 50ms

bfd modes

- 'async' send keepalives
- ...bog standard
- 'demand' out of band
- ...monitor traffic counters over the actual interface
- ...intimate knowledge of the dataplane counters
- ...if there isn't traffic within that timeframe, send a keepalive

specs can be stupid

`RFC 5881 - BFD for IPv4 and IPv6 (Single Hop)`

4. Encapsulation

BFD Control packets MUST be transmitted in UDP packets with destination port 3784, within an IPv4 or IPv6 packet. The **source port MUST be in the range 49152 through 65535**. The same UDP source port number MUST be used for all BFD Control packets associated with a particular session. **The source port number SHOULD be unique among all BFD sessions on the system**. If more than 16384 BFD sessions are simultaneously active, UDP source port numbers MAY be reused on multiple sessions, but **the number of distinct uses of the same UDP source port number SHOULD be minimized**. An implementation MAY use the UDP port source number to aid in demultiplexing incoming BFD Control packets, but **ultimately the mechanisms in [BFD] MUST be used to demultiplex incoming packets to the proper session**.

specs can be stupid

`RFC 5880 - Bidirectional Forwarding Detection (BFD)`

6.3. Demultiplexing and the Discriminator Fields

Note that it is permissible for a system to change its discriminator during a session without affecting the session state, since only that system uses its discriminator for demultiplexing purposes (by having the other system reflect it back). **The implications on an implementation for changing the discriminator value is outside the scope of this specification.**

specs can be stupid

`RFC 5880 - Bidirectional Forwarding Detection (BFD)`
4.4. Keyed SHA1 and Meticulous Keyed SHA1 Authentication Section Format
Sequence Number
The sequence number for this packet. For Keyed SHA1 Authentication, this value is incremented occasionally. **For Meticulous Keyed SHA1 Authentication, this value is incremented for each successive packet transmitted for a session.** This provides protection against replay attacks.

current status

- COMMITTED!
- ... kernel and userland
- ... enabled for a month, then disabled again
- ... still actively being hacked on

current status

- minimal implementation (all of the MUSTs)
- can successfully negotiate against many routers
- ...Juniper, Cisco, Brocade, Extreme Networks, etc
- ...uptime of at least 5 days on each
- basic logging
- route messages
- pf rules

current status

- moved to route
- ...we monitor nexthop, this makes sense
- difficult to adjust route UP/DOWN state for directly connected hosts
- ...punt for now
- special bfd flag (F)
- special route messages (RTM_BFD)
- magically supports multiple neighbors per interface

Simple setup

```
$ route -n change 203.0.113.9 -bfd

$ route -n show -inet
Destination  Gateway           Flags  Refs Iface
203.0.113.9  00:bd:39:6f:02:01 UHLcF  2    vio0
```

Simple setup

```
# route -n get 203.0.113.9 -bfd
   route to: 203.0.113.9
destination: 203.0.113.0
       mask: 255.255.255.0
  interface: vio1
 if address: 203.0.113.1
   priority: 4 (connected)
      flags: <UP,DONE,CLONING,CONNECTED>
        BFD: async state up remote up laststate down error 0
             diag none remote none
             discr 2258318855 remote 845809738
             uptime 04m08s last state time 08s
             mintx 1000000 minrx 1000000 minecho 0 mult 3
     use        mtu    expire
      20          0     9824
```

Simple setup

```
$ route -n monitor

got message of size 112 on Thu Sep 22 22:27:45 2016
RTM_BFD: bidirectional forwarding detection: len 112
mode async state up remotestate up laststate down error 0
localdiscr 3492152476 remotediscr 4117111943
localdiag none remotediag none
uptime 14s lastuptime 03s
mintx 1000000 minrx 1000000 minecho 0 multiplier 3
sockaddrs: DST
 203.0.113.9
```

Simple setup

```
cli> show bfd session extensive
                                          Detect   Transmit
Address         State  Interface          Time     Interval  Mult
203.0.113.1     Up     xe-0/0/0.0         3.000    1.000     3
 Client Static, TX interval 1.000, RX interval 1.000
 Session up time 5d 20:23, previous down time 00:01:21
 Local diagnostic CtlExpire, remote diagnostic None
 Remote state Up, version 1
 Min async interval 1.000, min slow interval 1.000
 Adaptive async TX interval 1.000, RX interval 1.000
 Local min TX 1.000, minimum RX interval 1.000, multiplier 3
 Remote min TX 1.000, min RX interval 1.000, multiplier 3
 Local discriminator 55, remote discriminator 4264428758
 Echo mode disabled/inactive Session ID: 0x101

1 sessions, 1 clients
Cumulative tx rate 1.0 pps, cumulative rx rate 1.0 pps
```

future plans

- actual manipulation of route UP/DOWN state
- "authentication" support
- Seamless-BFD (RFC 7880)
- multipath

future plans

- integrated knowledge in bgpd, ldpd(mpls), ospfd, eigrpd, etc
- switchd, vxlan, etc
- draft-ymbk-idr-rs-bfd

Questions?