

BFD

IETF 68

David Ward

Note Well

- **Any submission to the IETF intended by the Contributor for publication as all or part of an IETF Internet-Draft or RFC and any statement made within the context of an IETF activity is considered an "IETF Contribution". Such statements include oral statements in IETF sessions, as well as written and electronic communications made at any time or place, which are addressed to:**
 - the IETF plenary session,
 - any IETF working group or portion thereof,
 - the IESG, or any member thereof on behalf of the IESG,
 - the IAB or any member thereof on behalf of the IAB,
 - any IETF mailing list, including the IETF list itself, any working group or design team list, or any other list functioning under IETF auspices,
 - the RFC Editor or the Internet-Drafts function
- **All IETF Contributions are subject to the rules of [RFC 3978](#) and [RFC 3979](#). Statements made outside of an IETF session, mailing list or other function, that are clearly not intended to be input to an IETF activity, group or function, are not IETF Contributions in the context of this notice.**
- **Please consult [RFC 3978](#) for details.**

Agenda

- Chairs: Intro, administrivia - 3 mins
- Dave Ward: Changes to base and corollary specs - 11 mins
 - Doc status
- Dave Ward: New draft point-2-multipoint BFD - 13 mins

Presentations

- Base and corollary spec update

BFD Base Spec Changes

- Demand Mode
 - It was broken (could hang when coming up)
 - Now independent in each direction (multipoint uses it asymmetrically)
 - D bit now means "I'm up, you're up, don't send periodic packets"
 - Periodic transmission resumes if session goes down

BFD Base Spec Changes.2

- Session State Maintenance
 - Session state must be maintained as long as you receive packets from remote side, even when session is down
 - If no packets are received, state must be maintained for a detection time after session failure
- Gives systems control of remote transmission rate, even for down/admindown
- Allows systems to control session establishment rate

BFD Base Spec Changes.3

- Down/AdminDown Semantics
 - AdminDown means "down on purpose, please ignore"
 - Down means "forwarding path probably doesn't work"
 - Leveraged in Generic spec
- Receive pseudocode tweaked a little
- Some text was clarified or expanded based on implementor feedback
- Plenty of editorial corrections

Doc Status

- Base and corollary docs
 - Rev-06 has had extensive comments and review from providers and implementors
 - Multiple interoperable implementations known
 - 2 known complete implementations of all modes
- Go to WG LC (taken to mailing list, 2 week timeout):
 - draft-ietf-bfd-base-06.txt
 - draft-ietf-bfd-multihop-05.txt
 - draft-ietf-bfd-v4v6-1hop-06.txt
 - draft-ietf-bfd-generic-03.txt
- Rev needed to make compliant w/ changes:
 - draft-ietf-bfd-mpls-04.txt
 - draft-ietf-bfd-mib-03.txt

Presentations

- Point-2-multipoint

Multipoint BFD Overview

- Verifies connectivity of head->tail multipoint path
- Technology independent (IP mcast, MPLS P2MP, etc.)
- Does not verify tail->head return path
- Does not verify unicast head->tail path
- Optional notification to head of tail status
- Protocol timing/scalability driven entirely from head
- Runs next to Classic Unicast BFD
- Falls out of existing Unicast BFD spec (pretty much)

Service Definition

- Service defined as a base function plus a number of options
- Options may be enabled in any combination
- Base function: Unidirectional Transmission
 - Head sends periodic packets along MP tree
 - Tails detect BFD timeout, do "the right thing" (e.g. listen to another head)
 - Head ignorant of tails, no BFD packets sent to head
 - Simple, extremely scalable

Service Definition - option 1

- Option: Solicit Membership
 - Head sets Poll bit in MP transmission
 - Tails send unicast Final in reply
 - Tail transmission smeared across time specified by head
 - Head gets a Pretty Good idea of tails listening (unreliable)

Service Definition - option 2

- Option: Tails notify head of session failure
 - Head directs tails to send periodic packets to head when tail detects session failure
 - Upon session failure, tail sends bfd.DetectMult packets (smeared across time) and then quiesces
 - Semi-reliable (multiple packets are sent)

Service Definition - option 3

- Option: Verify Connectivity of Specific Tail
 - Head sends unicast Poll Sequence to specific tail (learned by solicitation or outside means)
 - Tail replies with Final (and without smear, so it's quick)
 - Head reliably learns tail state (if tail ever replies)

Service Definition - option 4

- Option: Some Tails are More Equal Than Others
 - Side effect of unicast Poll Sequence is that intervals carried therein override multipoint values
 - Head can thus raise transmission rate of individual tails for failure notification

Service Definition - option 5

- Option: Silent Tails
 - Tails may be provisioned to never reply to BFD even when head sends Polls
 - Allows for large numbers of second-class citizens in class-conscious tail population

Session Types

- Operation modeled as distinct session types:
 - PointToPoint: Classic BFD
 - MultipointHead: Session on head sending multipoint packets
 - MultipointClient: Optional session on head tracking individual tail
 - MultipointTail: Session on tail tracking head

Demultiplexing

- Multipoint (M) bit flags multipoint packets
- Packet demuxing rules select session
- Session type determines elements of procedure

Protocol Tricks and Hackery

- Multipoint packets all sent with Demand (D) bit set, tails cannot send periodic packets while session is up
- Required Min RX value set to zero means "no periodic transmission ever" (controls failure notification)
- Silent Tail = 1 means "no transmission ever" (no reply to polls)

Environmental Assumptions

- Tail needs to be able to differentiate between packets received on different MP trees if same head is going to be heard from on multiple trees
- Head is identified by source address (which cannot change)

Required changes to base spec

- Base Spec Changes
 - Add M bit
 - Define session types
 - Modify send/receive procedures for non-PointToPoint sessions

Next Steps

- LC all base and corollary docs
 - MPLS and MIB to be rev'ed
- Point to multipoint to WG doc immediately upon charter update
 - discuss on list
 - Edit
- By next IETF non-multipoint docs will be WG LC and marching forward
- No WG meeting at next IETF
- Very close to “done”