BFD IETF 105 Minutes - Montreal

Tuesday July 23, 17:10-18:10
Chairs: Jeffrey Haas, Reshad Rahman

Video recording: https://www.youtube.com/watch?v=0s3Qqm_a2RM
Audio recording: https://www.ietf.org/audio/ietf105/ietf105-sainte-catherine-20190723-1710.mp3

[Agenda]

[Chair slides - Jeff Haas]

Jeff Haas: BFD auth stalled on IPR. What is not there is text which says license-free or license-free pending action.
Mahesh: legal dept at Ciena looked at the template of the IPR and in their mind they follow middle of line text, language no stronger/weaker than other IPR in BFD WG
Greg Mirsky: another document used IPR similar to Ciena and was deemed potential hindrance to implementation. Disagree with lawyers statements. Most common form "no royalties if you don't go after me" (nuclear option), in legal terms "covenant".
Jeff: mutually not assertive is another term.
Jeff: we won't change the mind of the lawyers. What do we do as a WG? Lawyer pointed to 2 IPR disclosures on 5880 which didn't have the nuclear option. So there is precedent in BFD for that kind of IPR.
Greg: yes there is history but the question is whether the IPR is applicable to document. The update to extended BFD also has proposal for authentication. There will be IPR disclosure and the holder of that IPR follows "covenant".
Jeff: we'll take this to the mailing list for WG to decide publication/hold
Rakesh Gandhi: RFC4736 has similar issue and was published as informational
Greg: When mail goes out for the working group, the group should examine the IPR and decide whether it's applicable or not. If IPR is not applicable, I will support it.
Reshad Rahman: you're proposing that we look at IPR?
Greg: yes, each of us should make personal decision.
Martin Vigoureux: Speaking as RTG AD, agree with you that this is on an individual basis.
Greg: everyone must come to this conclusion themselves. We're not discussing applicability by IPR; that's why I'm not really saying what I believe in
Julien Meuric: IETF consensus won't determine what court's consensus would be.
Greg: I agree. You can ask someone with more professional IPR expertise. By reading their document, I think we can come to reasonable conclusion; stretch vs. close call?
David Black: +1 to Greg's comments. Nothing substitutes for your own technical judgement. Don't focus too much on what the lawyers say.

Jeff: For BFD large packets, Albert did padding at IP layer (FRR stack), tested against unmodified FRR stack and VMX stack, works as expected. We could go to last call.

[BFD service redundancy - Sami Boutros]

Acee Lindem: Single BFD session have different preempt status dependent on diag code?
Sami: BFD session for node liveness, not service
Acee: it's usually all or none. We already have this revertive/non-revertive in BFD.
Sami: no.
Acee: I'm thinking VRRP
Sam Aldrin: You're monitoring the link between the nodes; the BFD sessions. Interpreting BFD down as node failure? That may not be?
Sami: You're talking about split brain cases. You have a point. Both of them think the other is down. In that scenario, when it comes back up for non-revertive service, the other node can figure out whether it should bring it up on the other node or not or take it down (if the other side is up)
Sam: Assumption you're making assumption node failure is same as connectivity failure.
Sami: You're right. We can put some clarifications for split brain.
Greg: Only scenarios you need to cover is active and standby. There may not be more nodes that could provide the same service. E.g. 3 nodes, one is active, one is standby, one is other. Perhaps interesting scenarios, standby goes down and other node goes standby. Etc... Is point to point BFD is best mechanism? Perhaps look at BFD for multi-point networks. Acee mentions VRRP; that is p2mp. Works nicely for PIM-SM - multiple routers on shared segment. Active is at head end of session. DR/BDR or others are listening. In that, we might think on how to insert the idea of active is non-revertive. Take to the list and discuss if p2mp is more generic.
Sami: Option to consider. Diag code is simply to say whether I was up or not when I detected you went down

Santosh PK: bitmap comment for the WG. We are introducing new payloads for BFD. Should this be considered for BFDv2 or extended BFD? Wait for the discussion.

[Extended BFD - Greg Mirsky]

Motivation - what do we want to do with BFD? Monitor and quantify quality of BFD sessions. Loss of BFD packets could be used as crude performance monitoring mechanism. e.g. losing 2/3 packets. Performance measurement, path MTU, etc. Use existing mechanics. Must be backwards compatible and extensible.
Perhaps also work for intermittent authentication of BFD session (more lightweight)

Jeff Tantsura: Are you proposing to decouple liveliness from new capabilities?
Greg: Liveliness could be monitored by one session, another session might be brought up for extended. Poll/final sequence is a query (like ping). Can couple other stuff to the poll mode. So, they could coexist.
Jeff T - continue at 3.3 ms, but may do some things like delay measurement working on different timers e.g. 20secs?
Greg: Good question. May require additional specification of the capabilities.
Santosh: With this negotiation, if doing the poll and it doesn't support it, and it doesn't support does the lack of final bring down session?
Greg: Capability negotiation could be done at any stage. When session is up, L/D measurements could be done using poll sequence. Need to specify the frequency of the measurements as per Jeff's question.
Santosh: what happens when extended BFD is not supported, what happens? No reply.
Greg: if other guy doesn't support poll for extended, we should just do classical BFD
Santosh: in 5880, if we don't respond to poll, we drop session.
Greg: if extended poll doesn't go, we need to revert to classical BFD.
[authentication slide]
Greg: this capability is for the occasional authentication; e.g. poll. Idea is both sides advertise auth capability and they select largest/best.
Jeff: there are existing specs on how to do BFD auth
Greg: same way
Jeff: would you listen to auth header or capabilities?
Greg: auth header is based on config. In capabilities it's only for occasional authentication.
Rakesh: For TWAMP we had control channel singling, STAMP removes some of that. By adding complexity in BFD, I think we're going in the other direction.
Greg: It's for quality of BFD of session

[BFDv2 - Jeff Haas]

Jeff: we don't have to change the 5880 core BFD use-case. Should the BFD-like extensions be in IPPM WG?
Greg: auth of parallel session. Quality of bFD session?
Jeff: for quality of BFD session, isn't that more IPPM like?
Reshad: you could do that with IOAM.
Matthew Bocci: this discussion reminds me of MPLS-TP. we already had tools to do that, don't overload BFD. We used a separate control channel. Created MPLS performance monitoring framework. Nice in an MPLS context. Why do we replicate other mechanisms in BFD? We have these elsewhere.
Greg: w.r.t. perfmon it reuses 6374.
Matthew: but it's not 6374, we already have this.
Greg: 6374 only defines format, not transport. Issue with unidirectional measurement. Receiver gets metrics, what do you do? I like the 6374 PDUs. But it lacks some of the mechanics. BFD provides some nice mechanisms for those mechanics.   Measurements could be done using poll sequence, similar to STAMP.
Matthew: really using BFD plumbing?
Greg: reuse rather than reinvent.
Frank Brockners: negotiation and more flexible detection that are for perfmon, not entirely sure. One of the assets - BFD is acceptable because it's simple. If we're opening it up, does that hamper acceptance? I don't know. Figure out what we'd like to do and what existing tools we have. If we have parallel session doing TWAMP? IOAM? What do you get from these things. Once you open it up you may hamper BFD itself.
Greg: will note that IOAM is one way measurement. 6374 indicates that one way needs to be exported. similar to TWAMP when you do poll sequence. you could send back performance results.
Santosh: let's not open BFD format. But at the same time we need some extension to BFD. BFD stability draft - we wanted this feature dynamically but couldn't do in base BFD. Sami gave another example of light-weight parameters for BFD session. Shouldn't use it for LDM.
Greg: We can be conservative with what we put in.
Sam Aldrin: do we know what current BFD cannot do? Scope is so wide. Maybe we should define what it should do rather than it can do everything ; constrain the problem. Who is asking for this?

Greg: Several proposals for BFD that's not continuity check.

Sam: Need to do due diligence of existing tools. How does this play out with other existing mechanisms in other WG?

Greg: Santosh mentioned BFD stability. It was a BFD local feature. There was no mechanism for counting BFD packets were being dropped.

Rakesh: Current market direction is to simplify control plane. IPPM WG, the direction is to simplify or eliminate control signaling. TWAMP/STAMP fixed offset to make cheap for silicon. Message sizes are small, makes it easy for silicon to know how much to load. IPPM is the right place for this. Proposals to make BFD use TLVs is wrong direction.

Greg: I disagree.

Martin Vigoureux: BFD is a simple communication channel. Concerned that we're overloading BFD with info which is not related to OAM.

Jeff: BFD as protocol, not a given instance. Suggesting we don't want to overload core CC BFD use-case.

Martin: understood, but this is not OAM anymore.

Jeff: yes it's not the CC BFD has been doing.

Ketan: in favor of CC case. like this case. extensions that improve CC - great! Another session, etc. just to do the other stuff... before we go down that path, we should look at what the other problems are. See IPPM directions. if there are gaps, approach it that way rather than the other way.

Albert Fu: regarding BFD drops, features to see when packets are dropped and how to expose it would be helpful.

Jeff: cases where some extra data is needed, do we carry little more data for some cases. That may be where we draw the line. Longer term question is whether expanded charter or move to other WG such as IPPM for example.

Rakesh: Extend BFD as a control signaling mechanism for 6374 - probably not, don't think the right argument.

Jeff: will need longer conversation on alias.

Greg: MANET WG is interested in this, let's include them in the discussion.

Reshad: How many people read Greg's draft? (several)