DUANE WESSELS: Alright, everybody. Welcome back to day two of metrics. I think to start off this morning, first, we'll just sort of recap what we talked about yesterday, especially in the afternoon, because I'd forgotten all the things we talked about in the morning already. But in the afternoon, we went through three of the RSO metrics with the goal of getting some sense of agreement on thresholds for those metrics. And my recollection is that for the RSO availability metric, remember we spent some time looking at the formulas for parallel availability and whatnot, and we talked about how many root servers we considered needed to be available for the system as a whole to be up, and we settled on a formula which gave us the number eight.

We plugged that into the formula for parallel availability and came up with 96% availability for an individual RSO. So obviously, when we put this into the document, there'll be lots of text explaining how we got here and the rationale and that sort of thing.

BRAD VERD: Will the text include the amount of time and effort that went in? Like with this equation but saying that this equation kind of applies as a design method and not an availability type of thing, and then we kind of ...

DUANE WESSELS: Yeah, we can certainly put all that in.

BRAD VERD: I think it would be worth adding all that just so people know that this was – I definitely like the comparison to an airplane type of thing. That was very illustrative for people, I think.

DUANE WESSELS: Yeah, we can put all that in and talk about how this model makes certain simplifying assumptions that won't necessarily be true, but that's what we had to do in order to make progress and so on.

UNIDENTIFIED MALE: Just as a note, those simplifying assumptions actually I think only reduce the five nines to three nines, which is still perfectly reasonable for people. So that is that when two of those were on the same one, it looked like it only shifted it over one or two things in your table. But once we do the full test, we can ...

DUANE WESSELS: Yeah. So that was availability, then we moved on to RSO latency and we looked at the spreadsheet where people had been giving their suggestions, noting that these ranged from, say, 250 milliseconds to one second and we got everyone to agree that 250 milliseconds for UDP was reasonable and 500 milliseconds for TCP was reasonable. So we'll proceed with those numbers when we take this back to the caucus and the recurring discussions that we have via Zoom.

And then lastly, we talked about publication latency and settled on one hour as the threshold. We talked about why that is relating to the needs of the TLD managers and the way that the zone parameters work with

retry timing and all that kind of thing. So that's sort of where we ended up yesterday. Does anyone remember differently? Okay.

Now, the goal for this first session this morning is to talk about the correctness metrics in terms of the RSO metrics. So remember that there's currently two techniques defined in the document. There's one based on DNSSEC and one based on exact matching. We need to discuss, should we have both of these, should we continue with both of these, or reduce it down to just one?

And then as well with DNSSEC correctness, Paul Hoffman had some things that he wanted to discuss relating to maybe changing the way the method works a little bit and he had a proposal that I haven't had time to review yet for doing the random NXDOMAIN a little bit differently. I think that –

[PAUL HOFFMAN:] Or just not even using it.

DUANE WESSELS: Okay. Or maybe not even using it. So I think to start with, we should open up the discussion for, does it make sense to have both of these metrics in the document? Just as a reminder, we started with the DNSSEC validation-based correctness metric. My thought on that is that it was sort of a simple way to accomplish the correctness metric. The downside is that you don't get the verification of unsigned data, so that's really why the other one was added.

You could argue that the matching-based on can do everything that the validation one can do. Personally, I would kind of like to see an implementation of these and get some experience in seeing how hard they are to actually make work, but maybe that's not the best task.

So happy to open it up for discussion. I know Paul certainly has some opinions on this.

[PAUL HOFFMAN:] You've sort of given them.

DUANE WESSELS: But anyone else feel free to chime in.

BRAD VERD: Did anyone else read the caucus mail overnight? It was a message quite germane to this.

DUANE WESSELS: Are you talking about the discussion of the root server threats that's going on? Okay. Can you explain why that's germane?

BRAD VERD: I will. I'm just going to bring it up on the screen here. There's an RSSAC caucus message about treating DNSSEC in the root as a reliable indicator of correctness.

UNIDENTIFIED MALE: [inaudible].

UNIDENTIFIED MALE: [inaudible].

BRAD VERD: So I think we all know that there's not a huge amount of interoperability and completeness testing going on in the DNS validation market. It's kind of ad hoc and almost all of these developers on the open source side do a fair amount of testing because they live in a fishbowl and they don't want to be outed as the one who implemented it wrong.

However, the concern is raised that in many cases, wrong DNS information will not be detected, and in some cases, the mere act of changing the zone or maybe intercepting traffic and then forwarding it to the real nameserver will still give you an information leak possibility. I know that's not our concern, but I would like to know that the authors of the metrics work have read that mail from last night. It would only take two minutes. I can read it out loud if I can find it, or you can just find it and read it.

FRED BAKER: I'm looking at it, do you want me to read it? Or can everybody find it? I see [Kenneth's] nodding.

So I can do this, and this is from Michael Casadevall who's now a caucus member. He says – and I'm just quoting here, I'm not going to make any

commentary as it goes along. “Looking at the document, one thing that jumps out at me –“

BRAD VERD: Sorry, which document?

FRED BAKER: Oh, sorry, the thread started with Joe Abley talking about the new root ops document called threat mitigation for root server system. That was a thread that was started on the 30th. After that, Wes Hardaker and Mallory Knodel and Wes replied, and then Michael replied.

Michael says, “Looking at the document, one thing that jumps out of me is both data integrity and route hijacking. I feel like this document puts far too much faith in the validating resolvers to catch compromised root servers and route hijacking.

When I did work on creating a root server emulation setup for testing root IDNs, generally speaking, unless I went out of my way to force strict compliance, non-signed or invalidly singed data would just be accepted if it came from the root, even before I inserted my fake KSK into the mix.

The short version, if I can pop a root server and inject malicious traffic, I can't tamper with the root server data without causing signature validation failures. I can – emphasis added – freely change the NS records to MITM any DNS lookups to a TLD, especially if I'm echoing signed data. That’s a pretty effective way to essentially cast a dragnet of all lookups to a TLD.”

Paul, did you want me to go beyond that? That seemed to be the main point. Okay.

DUANE WESSELS: So I guess my thoughts on this as well, it's an interesting idea. I don't know if we should lend it so much credibility. I would like to see more evidence that this is really possible, I guess. And maybe Michael would be willing to do that work in the caucus, but at this point, I'm not sure it should change our thoughts and plans on the metrics work.

BRAD VERD: If having seen this we still feel that we should be making the choice that's been put on the menu about what kind a verification we're going to do in the metrics, then fine, just did not want that decision to be made in ignorance of this information.

UNIDENTIFIED MALE: Paul, when you read it, did you then come to a conclusion of whether we should be doing separate DNSSEC testing versus matching testing, or combine them? You obviously were more affected by it than I was. So what was your conclusion after reading it?

PAUL VIXIE: This guy doesn't sound like an idiot. I don't know him, but he sounds like he actually did something somewhere that produced the result he's describing. Under those conditions, I don't feel that the metrics document as currently drafted gives us a second set of eyes. So it may

be that we don't have zone MD yet, so there is no second set of eyes that we can ask for, but I think we should describe this as a potential gap and say that when some other additional verification of correctness is available such as zone MD, that will be added. That would be my only way to currently respond to the concern that was raised there.

DUANE WESSELS: Thanks, Paul. So we have talked a little bit about the zone MD concept in the work party and on the caucus mailing list. I don't think it solves the problem that we're trying to solve with this correctness matching metric, so the zone MD lets someone who's receiving the zone verify that they got the whole zone. I guess the only way we could use that is if the metrics platform was doing zone transfers from all the root servers. But it wouldn't tell you that a particular response was modified or not modified. So as one of the authors of the zone MD, I would certainly like to see it adopted and used in this context, but I'm not sure that it solves the correctness problem exactly. Russ.

RUSS MUNDY: Yeah, I've read some of the thread yesterday, and it sounds to me that having – this is at least a reasonable justification for having two ways of doing the verification, that we currently have [defined of a] second comparer is not as well defined as – the first one, the protocol defines it all, DNSSEC. So this is at least a somewhat additional view of what is correct data for the things that – why the design of DNSSEC were not included in the DNSSEC cryptographic verification.

[PAUL HOFFMAN:] I actually came to exactly the opposite conclusion because what Michael was talking about here is changing NS records which we can't test with the DNSSEC validation. The NS records was inf act my motivation for adding the exact match. And again, I believe exact match testing is a full superset of the DNSSEC validation which is why it replaced it.

My model was not the model that Michael had. My model was the fact that 80% of the resolvers out there don't do DNSSEC validation. So any answers they get – if we say, well, we can only validate certain kinds of requests like DS and such, that doesn't matter to them anyways, but the other one is that I think if we look at the attack surface, what would an RSO who was told, say, by their local government that they had to change something, the thing that they are most likely to change is the NS record, which is not what we were testing before.

So I wanted a way to do that, and Michael comes at it from a different aspect of it. If you change the NS record, even if it's going to a place it's going to give the right answers to somebody who can look at it. I agree with Duane, that's not my concern. My concern is the fact that so few resolvers are actually using the data we would want them to use for validation that we should be serving that 80% as well.

RUSS MUNDY: Well, and this was the reason for my conclusion, that having two methods is preferable because the one that's DNSSEC-based is cryptographically strong, well defined, thought through, etc. The one that's just a straight, if you will, bit compare, can be attacked itself. So you would have one method that gives you a stronger result of

correctness but didn't address every single piece of the content, and another one that addressed every single piece of the content but wasn't as strong of a verification of correctness.

[PAUL HOFFMAN:] So would you then be happy if the outcome of this was that there was just one correctness test and for any query in that correctness test, they could also be DNSSEC validated? That that had to be matched and DNSSEC validated, but for all the queries that did not have to be DNSSEC validated, it still would be acceptable to do just matching.

RUSS MUNDY: Well, I'm not sure if that's not just another way of saying we'll have two ways of doing comparison. But ...

[PAUL HOFFMAN:] Well, it sort of is, but I think what's attractive about that idea is that it's not published as two metrics. It's just one metric in the report. So then people would have to wonder, why is it 90% in one and 100% in the other? That would be confusing, I guess. So yeah, I think that's actually a good idea, is you could just add a little validation piece to the matching one to say if there's signatures, they must validate. If they don't validate, then there's nothing correct about the response. So maybe that's an easy way out of this.

RUSS MUNDY: Terry?

TERRY MANDERSON. Thank you for getting to the point where I really hoped you would get to, and that is just to have one metric that includes both properties of both matching and DNSSEC validation, keeping in mind the scope here is to ensure that the root servers have the right data and are serving the right data.

I'm not really going to entertain a huge amount – and this is personal opinion only – of discussion surrounding attacks on DNSSEC. Whilst they may occur, they have not yet been proven. They may be proven in the future because of insert any level of quantum computing. That's something pie in the sky at this point.

So I'd like to really ensure that this is really focused and well scoped to just the data that we distribute to the root servers. And let's not try and have a beef about the protocols itself.

ABDALMONEM GALILA: We could check the correctness of the data through every DNSSEC signed TLD publicly inside the zone. So [could check the public key with the DS record and the] zone file if this fingerprint which is the DS record is the same for the fingerprint of the public key, which will be, I think, solved for this.

LARS-JOHAN LIMAN: A scaling issue warning. That works of the root zone as it is today. If we move into the future when it's the [inaudible] zone and you have to go

to all of the children to fetch the public keys to compare them with the DS records, you have a bit of work ahead of you.

UNIDENTIFIED MALE: Sorry, I didn't understand that. If the query that we are testing is just the NS record for a TLD, why would we have to look at any of the children?

LARS-JOHAN LIMAN: Well, if you're going to validate the DS record in the parent, you will need to verify that that's a proper DS checksum of the public key, which isn't stored in the root zone but in the child zone. So you will have to go to the child zone to fetch the key to compare it.

UNIDENTIFIED MALE: I see.

UNIDENTIFIED MALE: I think that's encroaching on out of scope things because that's a data problem with the TLD. That's not the root zone serving incorrect data, that's the TLD manager having a mismatch between their key and their DS record.

LARS-JOHAN LIMAN: Yes, I fully agree and that's what I wanted to call out here, that this is one step beyond.

UNIDENTIFIED MALE: Although since we had been speaking yesterday, this isn't just about RSSAC 37 and such and there is a desire to have correct data in the root zone. At some point, somebody could do that. In fact, it could be RSSAC doing that. I wouldn't do it on a real-time basis, but I could certainly see a once a day going in and trawling to get all of them and such like that, to do that as a – is what we're serving what we think we're supposed to be serving?

LARS-JOHAN LIMAN: I would very much like to flip that around and say the TLD is welcome to check the DS record, because that will move – the scaling issue will be a totally different thing and you put the responsibility where it belongs.

BRAD VERD: We've had issues where TLD operators have not updated their DS records in their zone while they have been updated in the root zone and it turns into an emergency for someone else, not the TLD operator.

ABDALMONEM GALILA: Just a notice, most of TLDs don't have signature for NS record. They opt out of the signature. Maybe I am XYZ, I have 10 million domain names, I don't have signature for NS record. I only need signature for [A records or require A record.]

DANIEL MIGAULT: In this work party, we're interested in checking that the data from the root zone is appropriately served by the server. So what is the DNS data? Is it the data the root zone is authoritative for? In which case DNSSEC is probably sufficient to validate that it's correct.

DUANE WESSELS: Or some of the data, yes, but not all of the data will be validated [inaudible] DNSSEC.

LARS-JOHAN LIMAN: To me, the task at hand is to make sure that the entire content of the zone as we receive it is correct. And that includes a lot of records that are not signed. So the signature – DNSSEC as it stands, only zone is not sufficient. We need something more, and I like what you talked about before to compare, and if there is DNSSEC, to also validate.

DUANE WESSELS: Let me ask a slightly different question on this metric. The proposal in the document is that when you do correctness matching, you look back in time up to two days. So you store two days' worth of zone files that you could compare against. Is that the right amount of time?

One case in which it may not be the right amount of time is if you have a root server that's not able to receive updates for more than two days. It can still serve validateable data, but it would be potentially older than two days.

LARS-JOHAN LIMAN: My view on that is that it's definitely sufficient, because if we have that situation, we want a signal. We want the system to flag this. There could be a valid reason behind it, and that can be explained in the follow-up report or something, but ...

DUANE WESSELS: Okay. One of the reasons I ask is because when we got to the spreadsheet and everyone puts in their threshold for correctness, it's 100% across the board.

So if you've got an incident in a faraway place that doesn't get updated for two days, and if there's a probe that can hit it, it's going to fail the correctness and it's going to drop below 100%.

BRAD VERD: I agree with 100% correctness, however I also believe there'll be exceptions to every rule. However, the only legitimate scenario I can think of is something like the Arab Spring where a country disconnects itself and there's a conscious decision by root server operators to continue to serve the root while it becomes stale.

In that situation though, I don't believe any probe's going to be hitting it, much less reporting on it even if it is. And I think those are explainable. But anything outside of something like that, or I think the other scenario in which I know we've served stale data is when the earthquake hit Asia and cut a bunch of cables. There was a conscious decision to continue to serve data over there because if we turned it off,

it would only exacerbate an already ugly problem of sending more data over those cables as they tried to reach through somewhere else.

UNIDENTIFIED MALE: And you're okay with two days?

UNIDENTIFIED MALE: I think we need to define again a little more clearly what we mean by correctness, because we've talked about the zone correctness as a whole, but the tests that we're doing are really piecemeal pieces of that zone, it's not the full zone at any given point unless I'm misunderstanding these inquiries.

UNIDENTIFIED MALE: We're getting to that in the next question.

UNIDENTIFIED MALE: Okay. But let me just make sure all those will be addressed as we go along and I'm not missing something. So you asked for two days, and I'm thinking of a scenario where in day zero, you have three A records for reaching a TLD in the zone which is not signed but it's in the zone, sometime in the next SOA publication of root zone, [inaudible] and then it changes back to [three.] What is correct in that sense?

We've lost the fact that something changed, and maybe at that time, that was not correct. Are we just looking for a superset of the records

that we expect to see in the zone, or is it a more formal definition of consistency of those records with that SOA publication? So that's one.

And I think we are kind of conflating this notion of correctness with publication latency. We talked about staleness. And in my mind, that's publication latency problem, not a correct – I mean they are related and I'm not sure how do we keep the two metrics distinct.

UNIDENTIFIED MALE: Okay. So to your first question about the three to four to three, first of all, the way that the matching description is written is you get a DNS response, you would look for that data in any of the root zone files that you have in the last two days. So if your response has four, and if you can find a root zone at which there were those four address records, then it's a match. Does that make sense?

UNIDENTIFIED MALE: So suppose it was four and four was a problem, someone put in an extra record in some way and then changed to three and three. Is that a problem? Will we detect that?

UNIDENTIFIED MALE: I'm not sure what you mean by someone put in the ...

UNIDENTIFIED MALE: For the same reason that someone could take off an NS entry from the zone, someone could put in and enter a new A record, maybe before

observing any of the queries [inaudible] has a delegation or referral. So the addition of new unauthorized A glue record is also [inaudible] in my opinion.

UNIDENTIFIED MALE: So the wording in the current document might not be exact, but the idea is that whoever is checking this would check the answer gotten against the most recent root zone file. If they don't get a full answer, they go back another root zone file and back another, and then stop at two days. So if the man in the middle attacker you're concerned about is making an answer that was correct in a root zone file in the last t wo days, then yes, it will still pass.

But that would have been true, had they done – that already would have been true six hours ago or whatever when that root zone was there. So it's not the super set you were concerned with, it's an exact match against a root zone file and another exact match against a root zone file stopping two days.

UNIDENTIFIED MALE: So I think the right t erm to describe that is probably consistency, because we're looking at the consistency of the zone at a given point in time. The zone is not correct at the present time, it is stale. But it matched something that was correct two days ago, is my point.

DUANE WESSELS: This is not zone correctness, it's response correctness.

UNIDENTIFIED MALE: Okay. I don't want to go in the rat hole, just want to make sure that that difference is [caught.]

DUANE WESSELS: I want to harp on the two days just al b longer, because there was an incident back in August with C root – right, Paul? – where C root was I think exactly two days stale. So depending on timing, this could have impacted that correctness metric.

UNIDENTIFIED MALE: Yeah, we did not feel [we were in profile.]

DUANE WESSELS: By in profile, you mean what?

UNIDENTIFIED MALE: We did not feel that there was latitude for us to be that late and not having done the wrong thing. So that was an error and we have put various processes in place to make sure that it gets caught earlier and also can't happen the way it did this time.

DUANE WESSELS: Okay. But still, I think it's illustrative that it can happen unintentionally, and might happen again with somebody else. Alright, how are we doing on time? Go ahead, Russ.

RUSS MUNDY: Yeah, one of the points that came up – and I think it was Liman that made it – is that this check, this metric is supposed to have to examine that the response is consistent and correct with the appropriate SOA. And I was just going back looking at what's in there again and I'm not sure from – I think Paul's description, what we currently have is if there is a match in terms of the response, then it's a match and it's okay for any of them [inaudible].

But in the case of Paul's example of some government saying, "You must do X," and X ends up being somebody locally shoves in another NS record, do we have – I'm not sure that this would detect that.

UNIDENTIFIED MALE: Absolutely. And if it doesn't, it should. We need to fix it to do that.

RUSS MUNDY: Okay. That's [what I was reading it for.] So we need to look carefully.

UNIDENTIFIED MALE: [inaudible]. Absolutely.

RUSS MUNDY: Okay, good. Thanks.

UNIDENTIFIED MALE: I keep thinking that we're in a metrics work party not a zone validation work party, and so I thought the idea of the first DNSSEC check is just making sure we're getting a DNSSEC response and that it does validate, and I think that's good enough. And if we need another work party to go in and make sure that the entire zone and its contents are being served correctly, that's different than a metrics work party.

LARS-JOHAN LIMAN: Thank you. You just brought up a very interesting point. What if the root zone is distributed with an incorrect signature? Then it will not validate because it shouldn't validate. So it actually should pass.

UNIDENTIFIED MALE: [inaudible].

LARS-JOHAN LIMAN: If the root zone that I receive under appropriate forms comes with an incorrect signature, a signature that doesn't validate, then that's what I publish. And when you test it, it doesn't validate, because it shouldn't, because I'm publishing the zone that I got. So when that test fails, it should still give a green light for me because I'm doing the right thing. And that's [inaudible] point.

UNIDENTIFIED MALE: Thus my argument that exact match is all we need.

LARS-JOHAN LIMAN: Yes. I was just arriving at that conclusion.

FRED BAKER: Question. Is the correct response then that you should check the DNSSEC information as you download it?

UNIDENTIFIED MALE: Well, that's out of scope –

FRED BAKER: [inaudible] somehow it's corrupted.

UNIDENTIFIED MALE: I agree with that, however, there's no use to have a test or the ability to test something if you don't have some idea of what you'd like to occur if the test fails. So I believe that our choices are bad and worse, which is serve the bad information that we're trying to reject or serve the old information whose signatures are about to go stale.

FRED BAKER: Well, yeah, but if we're getting the wrong data from what appears to be the right [inaudible], then I have to ask why. Could it be that the route to the IANA was hijacked for example? But before I just willy-nilly started serving that incorrect zone, I would want to understand exactly what happened.

DUANE WESSELS: I think this is an interesting discussion but we're wandering off topic for the metrics work. Terry, you want to shut us down?

TERRY MANDERSON: I think you're pretty much close to what I'm about to say, is that I think we've gone into the realm of hypotheticals that probably won't apply. If I get a zone that won't validate, it won't be served. Simple. Nameserver won't let me serve it. My processes won't let me serve it.

So I appreciate the hypothetical, but let's please keep on track and on scope.

DUANE WESSELS: Yeah, I think this is important discussion to have, but not right now. So we can add it to an agenda later, I guess. Have we exhausted the question about two –

UNIDENTIFIED MALE: If people agree that we should just have one with two tests.

DUANE WESSELS: Okay. I feel like we've had a good discussion about that. Oh, Robert.

ROBERT STORY: I disagree that the matching is the only thing that's needed, because we've been promoting DNSSEC as the way to know that you're secure,

so we should be checking, if someone takes out DS records or signatures or whatnot, your test will still pass because the matching – no? Okay.

DUANE WESSELS: Maybe I should read the text in the section. Let me find it. Okay, so we're looking at section 5.1 in the document, and it describes for example that you do certain types of queries with 50% probability and other types of queries with 50% probability and they're all NS queries.

And then there's a list of rules for how you do the matching. If you get an RCODE=0 response, you make sure that the response, the AA [bit] is set. You make sure that the answer section is empty. You make sure that the authority section contains the entire NS [RR set] for the TLD that's present in one of the zone files. You make sure that the additional section records match those found in any of the zone files.

So the additional section is sort of optional. It could be truncated. You don't have to have all of them. You do have to have all of the NS records.

This currently doesn't say anything about DS because I think as proposed, the DNSSEC okay bit is not set, but we may modify that, and if we do, then we would add a check for the DS record. If there's a DS record, it has to exactly match the ones in the zone. And then there's a similar set of rules for if you get an NX domain response. You check that the AA bid is set, the answer section is empty and so on.

So I think as written, it's good, if not exactly the way it should be already. If there's something missing there – the intention is to do the

right thing and to check all the cases of possibilities of responses that you could get.

And I think what I heard from the group is that maybe what we do is that we combine these two or at least add DNSSEC to the second one and then drop the first one and call it good. Is everyone on board with that idea? So we end up with just one metric, we'll add DNSSEC checking to the correctness metric, and sort of be done with it.

UNIDENTIFIED MALE: Right. We would add [inaudible] for anything that we were testing that could be validated.

RUSS MUNDY: Yeah, and the one thing that I think we're faced with at that point is, are we going to maintain a 100% correctness requirement for the metric for everything? And I think if folks are agreeable that the matching correctness should also be 100%, then I think that challenge is taken care of, because everyone seems to be in full agreement that DNSSEC checking should always be 100% for correctness.

So are we okay with 100% for matching correctness?

DUANE WESSELS: Okay. So Paul, I think you had some proposals now to change things up a little bit even though ... Go ahead.

PAUL HOFFMAN: Well, yeah. Given a unified correctness test, I think a logical way to say, "Well, what are we going to check correctness on?" Is everything in the root zone, statistically over time, not every five minutes. Every record that is in the root zone that you would expect a correct answer for, which is every authoritative record. And then we need to – well, we don't need to, but I think it would also be good for us to do some negative checking, which is what you were doing with the random one.

So I don't think we can do this at mics right now, but my general proposal would be that the group figure out which is more important, like giving correct answers for positive things or correctly saying no for negative things and that might be 50-50, that might be 80-20, and it's easy, I think, to say, of the things that should be correct, is we actually check – randomly pick a record that's in a root zone and say we're going to check for that.

And then the question is, what do we ask for for negative ones? Currently, we have these made up TLDs, and we could ask either for – you were asking for DS, but we could ask for NS or whatever. Paul Vixie suggested yesterday that we actually check whether the root zone is giving an authoritative answer for a full query, because we don't get just TLD queries. So if someone says www.example.com and the root zone says "Here's the A record for www.example.com," if the root zone says, "Here's an authoritative answer" and it's to this address, we would like to note that.

So it's difficult and very rathole attractive to come up with what kind of negative questions to ask, and my proposal which I put in as a comment was that instead of doing obviously bogus ones, clearly identifiably

bogus ones, is we actually take maybe the top 1000 bad queries seen at root zones, like in DITL data, and do a statistical – ask for one of those.

So with the idea being those are going to be more dangerous if somebody is lying about those, and those are the ones that might be seen. We obviously can't do every A or AAAA query, much less a text record that could be dangerous or whatever, but those are the ones we're already seeing.

So that was my proposal for the negatives. I don't actually have an opinion on, should it be 50% positive 50% negative? I can believe almost any answer on that. But for the negative queries, I think the most popular ones that we're already seeing now should be the ones that we check for answers on.

FRED BAKER: I like your DITL idea, pulling of the 1000 screwball [names that we find.] I have been worrying lately about a set of things that I find. RFC 8483 says that the yeti project resigns the entire zone. I've seen three articles that say that certain DOH servers are being used to manage botnets, and people are uploading text records and their stuff there. And I find myself thinking that those categories of errors are sort of what we're looking for in this metric. Is the guy operating in a bad way?

So I wonder, is there a way that we can use this to detect for example the uploading of text records to manage a botnet? Is there a way that we can use this to detect a service that is incorrect?

DUANE WESSELS: My understanding of the text records via DOH stuff is that those are being put into large public DNS provider caches, right?

FRED BAKER: That's correct. Yeah, they're going into whatever – well, Google DNS is the one that's being used. Could be any of them, but that happens to be the one they're using. I don't have it out for Google, but I worry about that.

DUANE WESSELS: Yeah, so none of the metrics stuff that we're proposing would send queries to Google DNS, so I don't see how we would have any visibility –

FRED BAKER: It could. The obvious next step is to test places that people are going to for their DNS data.

DUANE WESSELS: Sure. It could be done. To me, that's out of scope for what the work party was asked to do, so it feels like new work to me. I'm not sure. I don't see that fitting here. Paul?

PAUL HOFFMAN: I agree it's out of scope here but I would like to go further. Given that this is a hyperlocal root service option, there r a fair number of things that Google's DNS servers or other recursive nameservers could end up talking to that aren't us, and failures that we could measure by talking

to some recursive nameserver may have nothing to do with any of us due to the fact that root name service is not exclusively provided by us.

We also have to look at the upcoming trend of resolverless DNS where the DNS binding's necessary to fetch objects referenced in a webpage are going to be included in meta headers in the webpage so that there will not be any DNS queries at all, and yet the DNS information will be consumed without verification of any kind by a number of browsers.

So I think that anything that has to do with looking at how the world experiences DNS is permanently out of scope for us who are concerned only with publishing some part of it.

[KEN RENARD:] Paul, getting back to your question a minute ago, the attacks mentioned here, the government taking over, saying you must respond this way, I think some of the attacks we've seen this year really geared towards existing domains versus just random things. If I can redirect a country code or something like that.

So that to me would skew things towards checking existing versus nonexistent – I'm not going to offer any numbers though.

DUANE WESSELS: A couple questions, Paul, about what you proposed. Using the DITL data, I think, is interesting. One problem of course is that by using the DITL data, you sort of create a feedback situation where you amplify those queries and they continue to float to the top of the list. In the [inaudible] [annex D] proposal in this document, there's a query name

which is prefixed with a known string. The idea there was that you could pick out that string in collections like DITL.

Yesterday, Paul Vixie argued for sort of the opposite, that the sort of queries that we make should maybe not be so easily identifiable so that they can't be seen and singled out. What do you think on this? Should these be identifiable as probe queries, or should they just look like normal traffic?

PAUL HOFFMAN: I believe our description of them would be more reasonable if we said these are real live ones. If we go towards “They should be identifiable,” then what you’ve done is perfectly fine, that it is easily identifiable, separable.

My logic here was more if someone’s looking at this, they're going to say, well, no one’s ever going to ask for those, so no sane rogue instance would have ever changed them. All you’re testing for is if a rogue instance is giving out generic addresses. So that’s like one query that you could ask.

I can't believe that an instance would answer some of them one way and some of them another. I can believe that a rogue instance could do – no offence – site finder equivalent, but that would be then universal.

So that’s why I went with my way, but I totally agree that there are side effects of doing that swell. And especially if we use the DITL data, it’s going to have to be clear to everybody what these queries are, and if

someone's rogue and they said, "Oh, but we don't want to do these," they could do it.

DUANE WESSELS: So one of the reasons that the [inaudible] [annex D] stuff was defined the way it was was – in the context of DNSSEC, I think it lets you identify cases where a TLD has been inserted into a certain range. So you maybe wouldn't know what that new TLD was, but you'd be able to detect that something had been inserted in this range where there should have been nothing. But yeah, I agree that that particular query is not going to occur naturally and maybe not something that would be answered for.

RUSS MUNDY: I'd like to hear some additional views from the work party about, should we be essentially doing kind of what I would call synthetic queries that are identifiable, that are clearly associated with this system? Or should they be just like ordinary queries? We've only heard from a couple of people. What do others think about that? Real or synthetic? Which is better?

Yeah, Duane just said, why not do both? And that I think is a possibility. Daniel.

DANIEL MIGAULT: The only concern I have regarding – so the good aspect of [reusing] data, it's actually data that's being [seen] on the network. What I'm trying to think about is that if you remember sometime we have very popular TLDs that are not TLDs, and you had a huge amount of that

traffic and some people were treating those specific data in a special way.

I'm just a little bit skeptical that – would the request actually reach the server? Or maybe it's going to be dropped before. Some ISP might have some mitigations in place for some queries. So [they might be acted in] man in the middle, but I'm not sure it really makes sense that the thing is running in my head and I don't know what to think about.

DUANE WESSELS: That is, I suppose, a possibility, but I think our intention with the vantage points is that they should certainly be located on networks where there is no man in the middle-ing going on. So if such were detected, then that vantage point would have to be relocated or shut down or something. It's not a reliable vantage point anymore. But yeah, detecting something like that I guess could be done via these measurements.

Well, as you said, Paul, I think at the mics is not the right time to fully design this. So we need to work offline with the work party and come up with some more specifics to look at.

PAUL HOFFMAN: Both on the split, if there is a split, and also of the negatives and the quality of the negatives.

DUANE WESSELS: You're right. And we could also turn to the DITL data to find what the distribution should be, whether it's, like you said, 80-20 or 25-75, whatever the current split of traffic is for ...

PAUL HOFFMAN: Oh, I'm sorry, I was not suggesting that.

DUANE WESSELS: I am suggesting that.

PAUL HOFFMAN: Oh. That would be all negative, basically.

DUANE WESSELS: Not necessarily.

PAUL HOFFMAN: Not if you take out the Google Chrome queries, but yeah.

DUANE WESSELS: So absent any other opinions or indications, we can use real traffic to find out what's an appropriate split of existing versus non-existing names.

DANIEL MIGAULT: the other concern I have reusing DITL data is that we're moving from something that is – we're making the measurement more complex with centralization, with related to other data, and it might also – I don't see how it's going to evolve over time, but the more complex we make the system, the less scalable it's going to be. And that's my concern. Typically if we need real-time synchronization and all these kind of things, it's impossible that we would be in a long-term may be able to expand that to regular probes, or this kind of thing. So that should be sought in the design, I think.

PAUL HOFFMAN: I think this would all be good for on the list, because I think as we hear each other's comments, I just all of a sudden had an "Oh" from something that Duane said five minutes ago. And I might actually just drop my whole proposal altogether. So yeah, let's do this on the list.

DUANE WESSELS: Okay. Alright.

PAUL HOFFMAN: It's NSEC, not NSEC3.

DUANE WESSELS: Okay. That sounds good. What is our break time? 10:15?

PAUL HOFFMAN: 10:30.

DUANE WESSELS: 10:30? Oh my gosh. Okay. Let me go to my notes. I think we've exhausted the discussion on correctness. If everyone agrees, then we can move on to another topic in the last half hour of this session. Does that sound fine? Now is your chance to say something about correctness. Okay.

I want to talk about RSS availability, and I'll ask Ozan to put up that slide deck from yesterday. There was one slide that we didn't get to. We were talking yesterday about this K [inaudible] and we actually arrived at some consensus that we should K equal to 8. Can you go to slide 14, Ozan?

This is a proposal that I thought of. This is not in the work party document yet, but this would be a new way of measuring the availability of the root server system. Using the formulas, the expected availability is, I think it was, [five 9s]. But this is a way to actually measure the availability now that we have a value of K.

So we've done step one. We've agreed on what K is. In step two for each measurement interval and for each vantage point, you do the measurements and you calculate the number of root servers that respond to the query. Each vantage point hits all the root servers and calculates how many it got responses from. Then in that interval if the measured number of K is greater than or equal to 8, then we say that the root server system was available from that vantage point at that time.

Then you can aggregate all those. So this would be a Boolean measurement. Either it's available or not available from that vantage point at that time. Then you aggregate all those to calculate – here it says daily. Maybe we end up with monthly, but you would calculate a daily root server system available as the number of vantage points that had availability divided by the total number of vantage point measurements. Again, we expect this to be [five 9s], so we expect this to be 100 percent. That's our threshold.

[RUSS MUNDY]: Reading the title there, [or assess] availability as a system from that vantage point, to me, could read as K=1. We're saying K=8. So at least 8 RSOs are available. That's what we're actually measuring. One could think of it differently as a K=1. But if we're designing for eight available, that's what we're saying. If only seven are available, then the RSS is down. A little bit misleading maybe.

DUANE WESSELS: Misleading how?

[RUSS MUNDY]: Misleading that if you have this Boolean that says RSS is not available right now, people are still getting answers. It's not what we want. It's not good, but maybe that's a marketing problem as well.

DUANE WESSELS: Here's my struggle. If we're going to have a root server system availability as a metric, we need a way to measure it, right?

[RUSS MUNDY]: Right.

DUANE WESSELS: So I'm trying to get us there, and it seems like we had agreement on the idea that you need eight root servers to be up at any given time for the system, for the RSS, to be available.

UNIDENTIFIED MALE: Yeah, I struggle with this too. Sorry to interrupt. It's not we need eight to be up for it to be available. We need eight to be up to meet the 100% availability that we all agreed on. Does that make sense? So it's not black and white, like he says. Is there a way to respond where if you drop below eight, it's not that it's not available? It's just that we're below our expected threshold.

DUANE WESSELS: Let me go through it again just in case we're not on the same page. There's a vantage point and if that vantage point in some five-minute interval can only reach seven, that vantage point at that time is – it says, for me the RSS was not available. So that is one data point out of 5,760 for the day.

UNIDENTIFIED MALE: Can you go back up to Slide 9? If you're over at seven, you're [four 9s], [not five 9s], right?

DUANE WESSELS: This is predicted availability.

UNIDENTIFIED MALE: Right.

DUANE WESSELS: This is a formula that can predict, but we need a way to actually measure.

UNIDENTIFIED MALE: So if one of those 20 vantage points didn't get a response, then it failed. But if they all got responses – if one server was up that all 20 vantage points got a response, then the system is up, right? It's available.

UNIDENTIFIED MALE: Yes, it is.

DUANE WESSELS: Okay. So we have 20 vantage points. They can all reach 1 out of 13, and you want to say that at that time the root server system was available?

UNIDENTIFIED MALE: Mm-hmm.

UNIDENTIFIED MALE: [For a metric].

DUANE WESSELS: So the only time that the root server system is not available is when a vantage point can reach no root servers?

UNIDENTIFIED MALE: Is this really saying whether or not the root system is available or whether or not we're able to reach the number of letters that we think we should? Just the same way when your latency is one and a half seconds, you're going to go red. That doesn't mean that you're not reachable. It just means things aren't good enough. So I don't think the availability going red is saying it's not available. It's not as available. We don't have as much confidence in the availability as we want to.

UNIDENTIFIED MALE: Sure.

RUSS MUNDY: Do we need to rename this metric from an availability metric to something else? Because what I've heard the discussion here in the last few minutes, it sounds like people are saying we are trying to define that if you get any response from anyone that [inaudible] the system, then that's available. Personally, that's been my view for a while. Brad?

BRAD VERD: I think what people are saying is if you receive queries from less than eight, we are not not available. Does that make sense? Meaning greater than zero and less than….

TOM MIGLIN: I have a question sort of related. If it fails this metric, what do want to have happen? Is it just recorded and used for, I think what we say in there is, long-term RSS behavior? Does anything happen? Do we have any action to be taken if we fail this? So what's the purpose of [the metric]?

DUANE WESSELS: Yeah, you were not here yesterday, is that correct?

TOM MIGLIN: No, [inaudible].

DUANE WESSELS: We talked about this a little bit yesterday.

TOM MIGLIN: Okay.

DUANE WESSELS: In the context of the RSS thresholds, there probably is – well, it kind of depends, I guess, on whether an action gets taken. It depends on the [governance] model at the time and so on. In some sense, you can consider it informational. RSSAC may want to talk about why is this happening, or the root server operators may want to talk about why is this happening. But there is maybe no direct action that can be taken at that time.

UNIDENTIFIED MALE: Just to expand on that, the action would be taken from this group and/or the root operators. It's to inform them.

BRAD VERD: Okay, so then when do we want the action? What threshold do we want the action to be taken? So maybe rename the metric to make it mean more what we want it to say, what we want to do, like Russ was saying.

UNIDENTIFIED MALE: I think I like the proposal on 14 because I think we are maybe losing perspective on who this metric is for. On one side from the publication point I completely agree that even if we are able to provide one answer form the full system, we're good. But I think what we arrived at yesterday was that if we have fewer than eight, there is a potential for someone who is depending on that system to not be able to resolve. So I think this proposal captures the fact that there are some who are dependent on the system who will be affected, and I like the proposal.

UNIDENTIFIED MALE: [I'm with Russ that] maybe we should find a better name for this. Redundancy level or something. Or we need to have a scale rather than black and white.

DUANE WESSELS: I'm struggling a little bit with the idea of renaming it.

UNIDENTIFIED MALE: Okay, it's not a strong opinion.

UNIDENTIFIED MALE: Can you go back to your term of saying it's unavailable and trying to explain that? If it's below eight, you're saying the root server system is unavailable.

DUANE WESSELS: If for a given vantage point, if it's below eight, that vantage point at that time the root server system was unavailable.

UNIDENTIFIED MALE: You can't say it was not 100% available? It was [four 9s] available.

DUANE WESSELS: But all the vantage points get aggregated into a broader daily average or metric, if you will. Let's take a simple example. Let's say the measurement system works perfectly. We have 20 vantage points. We get 5,760 measurements a day. In one day, if one vantage point at one

time has less than eight, that's 5,760 minus 1. So the overall availability for that day is whatever that is: 5,759 divided by 5,760. It's 99.99-something. So that day, our availability was just under 100%.

UNIDENTIFIED MALE: I think Karl had a….

DUANE WESSELS: By the way, my goal is to come up with a way to measure RSS availability. If someone has a different approach that they think is logical and reasonable, let's talk about it. But this is what I came up with.

RUSS MUNDY: Duane, I like the description. The concern that I have and the reason I suggested maybe we try to find a new name for it, to a great extent is a result of our friend Paul Vixie's comments earlier, is Paul gets to talk to the press probably more than any of the rest of us. So for that privilege he gets to know what they say and do about what he says and how they rearrange the things that get said. There's, I think, a certain amount of attention that's going to be caught somewhere in the press when we publish this. And the term RSS availability, the inverse, the non-availability, has a very negative result.

So that's why in terms of coming up with a more effective descriptor for it because it's not that a vantage point wasn't getting a response. It was not getting as many responses as we would like it to get by the definition here. And I think it's a perfectly fine approach in definition. So that was why I was thinking having a more neutral, if you will, less

inflammatory way because you get the press saying, "Oh, the root server system is not available blah, blah."

UNIDENTIFIED MALE: Yeah, it seems like this – I'm kind of with Brad in his thinking. This is sort of a pass/fail. Is the root server system 100% available? If it's not 100% available, it falls into a [degraded] state.

UNIDENTIFIED MALE: Okay, so maybe. What's up on the screen here is a proposal for a way to calculate availability. The threshold that we set is another question. We could say, I don't think we would want to say, but we could say given this metric our threshold is 99%. Now that has the marketing problems that we talked about yesterday. So really you want to say it's [five 9s], it's essentially 100%. But it doesn't have to be, right?

UNIDENTIFIED MALE: I think the problem is that availability means you are able to reach at least one.

BRAD VERD: If I may really quickly. I think I can solve this. Sorry. I've been wracking my brain on this. I think this is a terminology issue. I think your math is right, and I like this. I think this is good. But I think we don't say that if you don't get an answer, it's unavailable. I think if you don't get an answer – and you don't change your math – you get an answer and it's 5,759, but the root server threshold has been breached.

UNIDENTIFIED MALE: Performance [inaudible].

BRAD VERD: Performance or something. Performance threshold has been breached, versus saying it's unavailable. It's just a messaging problem, that's all.

DUANE WESSELS: And is this a messaging problem for the RSS but not for the RSO metric? Because we spent all day yesterday talking about RSO availability, and this didn't come up.

BRAD VERD: I don't remember an example where it was pass/fail for availability of the RSOs. Maybe it was and I missed it.

DUANE WESSELS: Again, the output of this is a number.

UNIDENTIFIED MALE: [inaudible]

DUANE WESSELS: The comparison of that number to the threshold is a pass/fail, but the output of this is a number. Just like in the RSO case, you send inquiries.

You calculate how many responses you got. That gives you a number. You compare it to a threshold. So there's some similarities here.

BRAD VERD: I think we're in agreement basically saying this math is correct and you come up with a number just like you would with 5,759. It's just not that that vantage point is not saying that the root server is unavailable. It's saying that the performance threshold has been breached. The performance threshold of 100% or [five 9s], whatever it is.

UNIDENTIFIED MALE: So I agree with Brad. It's probably wording. But what I would propose is that you introduce the K availability and K would be larger than one. It would be eight availability, and K=8. Then you don't get confused [with] having less than 100%. [You're going to have 800%.]

UNIDENTIFIED MALE: 800%?

UNIDENTIFIED MALE: Well, if you can reach 8 servers, you're 800% available.

UNIDENTIFIED MALE: Two comments. Why don't we call that redundancy level? B, having one unavailable, I think that's okay because then that takes down our redundancy level from 13 to 12. And we've already stated in a different document that we can live without one of them, so it is actually okay.

We can decide that we have an optics problem with that as well, but I think we have less of an optics problem with the individual measurement for the root server operator as opposed to this RSS metric.

DUANE WESSELS: So I want to make sure I understand. You would like this metric to not have the word availability in it but to have the word redundancy in it keeping this math?

UNIDENTIFIED MALE: Oh, the math is fine, absolutely. Yes.

DUANE WESSELS: So the output of this would be RSS redundancy equals 99.9%? Or RSS redundancy equals 100%?

UNIDENTIFIED MALE: As long as all the probes reach more than eight, the redundancy level is 100%. Or you can even put it a higher number, but that will look strange. But as soon as one of them goes below, so you no longer have 100%, your numbers there are going to start to show nines. That's your redundancy level.

DUANE WESSELS: So if I told you my car is 100% redundant, what does that mean to you?

UNIDENTIFIED MALE: I have no idea because I haven't seen the math for your car.

UNIDENTIFIED MALE: Duane?

DUANE WESSELS: Go ahead, yeah.

UNIDENTIFIED MALE: I wanted to answer your question about yesterday. We kept using the word availability even after Paul showed us that we were doing the 8 of 12, and I think that was a mistake. I think that in fact we at the end of the day, again using similar math, we came up with a way of predictability. But I don't think that was availability because, as we admitted, the first one might have failed and we needed to have 8 out of 12 be available to increase the chance it wouldn't.

I would say either performance or redundancy would also apply to where we ended up with yesterday. So even though the measurements are of availability, the metric that we came up with – I'm sorry, yeah, we changed the metric to use your math. The metric is not availability. It uses availability measurements, but the metric was redundancy or performance.

DUANE WESSELS: You're talking about RSO? The RSO case?

UNIDENTIFIED MALE: Yeah, that's what I meant by yesterday.

DUANE WESSELS: Ozan, can you put the slides on Number 2 or something like that? Number 3. This is a terse version of what the document says for RSO available. It's very simple. You send queries, count responses, and you divide them.

UNIDENTIFIED MALE: Yep.

DUANE WESSELS: So everyone agrees that we can call this availability?

UNIDENTIFIED MALE: Mm-hmm.

DUANE WESSELS: This is not redundancy, right?

UNIDENTIFIED MALE: Right.

DUANE WESSELS: Okay. And the long discussion we had yesterday was what do we want our threshold for availability to be. And we used the formula and we landed on 96%.

UNIDENTIFIED MALE: Yes.

DUANE WESSELS: Okay.

KENNETH RENARD: [inaudible] redundancy to [work back to] availability.

DUANE WESSELS: Okay, just so I understand there's no proposal to change anything about this. We're good with this, right? Okay.

UNIDENTIFIED MALE: I'm not convinced we're good with the name of the threshold.

DUANE WESSELS: For this one?

UNIDENTIFIED MALE: For this one, correct. I believe the threshold we came up with, as Ken just said, because we worked back for the needed redundancy in order

to get [five 9s] or [four 9] reliability. At that point, it is no longer availability, I believe. I believe that it is either performance or redundancy. So the threshold that we got is based on availability measurements, but it is not an availability threshold.

UNIDENTIFIED MALE: For the RSO.

UNIDENTIFIED MALE: For the RSO.

DUANE WESSELS: I don't see how you can have different names for these things. This is how you measure RSO availability.

UNIDENTIFIED MALE: Yes.

DUANE WESSELS: And then you have threshold for it.

UNIDENTIFIED MALE: That's not the threshold we ended up with.

UNIDENTIFIED MALE: I agree with Duane here.

UNIDENTIFIED MALE: Okay.

UNIDENTIFIED MALE: We have a metric here for the availability of a single RSO, and by combining the RSOs we get a redundancy measure for the system and that has a different threshold. So the 96% is an availability threshold for the RSO. And by using those numbers in combination with the new number, the redundancy test for the system has to rely on the availability and also the number of RSOs that can be reached. Because that's another parameter that goes into the redundancy equation.

UNIDENTIFIED MALE: But it's a parameter that we used to get the 96%, which is why I'm saying.

UNIDENTIFIED MALE: That's one we decided to have to find out what we needed. So we did a backward math to set a level. But then we can do [count] numbers to see if we reached that level.

UNIDENTIFIED MALE: [inaudible]. Okay, I can see that, yeah.

[DANIEL MIGAULT]: Duane, why don't we simply sum the availability provided by the RSO and then compare it to the thresholds? Because if we have 8 ways or 12 or 13 ways to get the answer, well it's more than 100%. It's 1200%.

DUANE WESSELS: So what Daniel is proposing is you've he's got this metric. RSO availability. Let's say, in most cases, it's 99% for all the RSOs, so Daniel's proposal as I understand it is you calculate RSS availability – or whatever name you give it – as the sum of these, so it would be 13 times 99, you would get almost 1300% as your RSS availability. Then you would say your threshold at 800%.

UNIDENTIFIED MALE: So that's almost your K number.

DUANE WESSELS: Right. I can kind of see it. It's a little bit strange to me because I don't generally like percents that go above 100%. It's kind of a weird concept but –

UNIDENTIFIED MALE: It says redundancy.

UNIDENTIFIED MALE: Yeah.

UNIDENTIFIED MALE: I have a challenge with redundancy but I have a question, Duane, in your other equation on availability for the RSS. I hate throwing out the exceptions. What I'm trying to understand how you would report … say there was a really bad day and bad things happen and it happened for more than 24 hours, and we were only 95% available, meaning you've got answers from 7 but not 8. In that math, would it be you'd be zero percent available? Or would it be 96% available as you got answers from 7?

DUANE WESSELS: Well, it would be 96%, 95% available. No. Okay, we're talking about [inaudible] all day?

UNIDENTIFIED MALE: All day.

DUANE WESSELS: It would be zero percent.

BRAD VERD: How do we account for that? That's the issue that I'm trying to account for. Because you're available, you're just not 100% available or [5, 9] is available. You're 90-something available because you're getting answers from 7, just not 8.

UNIDENTIFIED MALE: Then Daniel's math works just fine.

BRAD VERD: Okay. Except – I don't know – it goes against a lot of –

DUANE WESSELS: Right. That is true. If that were to happen, if there was an all-day attack that took out more than … left people with less than 8. Then for that day, by my formula you would say the RSS was zero … or you would say it didn't meet the threshold. But in fact, it would be zero percent, right?

BRAD VERD: I'm okay if you say it didn't meet the threshold but you weren't reporting a metric that said you were zero percent available because you are available, it's just it's something above zero but below the threshold. If there was math that we could figure out what that is then I think you're in a perfect world. We just don't know what that is yet, right?

DUANE WESSELS: Well, we're trying to use as much math as possible here to get us this metric. If we go through this exercise with K = 8 [inaudible] and we say, "I don't think so," then maybe K = 8 is not the right number. Maybe it's 3, maybe it's 1. I don't know. But what I'm trying to get us to is a way to measure this and report it.

LARS-JOHAN LIMAN: I'm trying to wrap my head around the different mathematical models because I think they have different properties and we probably want to know … If you look at Daniel's model, and I kind of like it, if the entire system is out for one hour and then all of this works for the rest of the day, we probably still have a number above 800% so it would be working. But I would like this metric to somehow signal that that wasn't a good day.

DUANE WESSELS: Yeah. That would be one of the disadvantages of what Daniel proposes. You could be down for a large fraction of the day and still be above 800% or the month or whatever it is.

LARS-JOHAN LIMAN: Right. I also don't like the black and white that if we don't have 100% availability, we suddenly have zero. To me, it's not black and white. So the threshold part is very [inaudible] to me. We have a threshold and we're below the threshold. That's something, and that should be flagged with a red flag. Say we actually broke the threshold, but it shouldn't say zero and unavailable because I think that's an optical problem that we don't want to have.

DUANE WESSELS: Okay. Right. To be clear, the procedure that's proposed on slide 14 gives you a number. It's not zero or 100. It gives you a number, which most days we would have [set] to be 100%. On a very bad attack day, yeah, maybe zero, maybe 50%, I don't know. It depends on the nature of it.

BRAD VERD: It seems like the other equation is maybe not complete. What I mean by that is it's the right math on the right path. Instead of stopping there, maybe we need to somewhere in there – your 288 times 20 is 288 times 20 times the number of responses if below 8. It's times 8 or times however many responses you get equal or less than 8. Does that make sense? So that you don't end up with a zero number but you end up with a percentage somewhere. And I don't know what the math is to make that work, but if I'm getting 20 vantage points, [280] measurements, and I'm only getting 7 responses, what percentage of that is that if I got … versus 8. Does that make sense? So maybe it's divided by whatever that number is times 8. Then you end up with an actual percentage. I'm not writing this out but –

DUANE WESSELS: We have to whiteboard or something.

BRAD VERD: Yeah. We'd have to whiteboard this out.

DUANE WESSELS: I hope that we all agree that in some time interval, to say that the root server system was meeting our expectations of performance, that there's some minimum number of servers that a vantage point should be able to reach. So if 8 is not the right number, if we're nervous that 8 is too high then maybe we should consider –

BRAD VERD: I think 8 is the right number. I think reporting zero is … I mean we need to figure out how that number gets reported to 95% or something. Is that possible? Can we do that? Or is that just not –

DUANE WESSELS: I guess there's a couple of options. Yesterday we talked a little bit about how … okay, sorry, Robert. Go ahead.

ROBERT STORY: On the mailing list and since the beginning of the Metrics Work Party, there's been discussions where there had been people – including me – arguing that RSO availability is positive if everybody can get a response from anyone. So, I get the feeling that we also want this idea of one really is a good, healthy number, and I'm thinking maybe we need two different metrics. One actual availability with everybody's getting a response, so we just need one, that's fine. And then a second for the redundancy if we fall below 8, we should start paying attention, something bad is happening. The system is still available but this other metric, the threshold about redundancy means that it's not great.

DUANE WESSELS: I hear you arguing for K = 1 as the minimum performance level –

ROBERT STORY: For availability.

DUANE WESSELS: For availability.

RUSS MUNDY: One of the things that I think has been causing it. Some folks' concern is how much of what we've been discussing is going to be information that's put out publicly and how much is going to be available to the RSO, the operators. And I think early on, we agreed that for each RSO, they could get all of their information. I don't know that we've talked about it for the system as a whole but it would seem, for a system as a whole, it would make sense to be able to make that same statement but there's never been an [attempt] to publish all the details from this. So I think we need to reach as a point of agreement of how we describe what is going to be publicly defined and published on an ongoing basis. I think that's what we need to focus on getting to here. This approach looks like a real good set of math to handle that. It can probably be applied to our four different transports and work that way. Then if K is 8 then we more or less solved what we've said it's going to be and then used, and if that is a defined as either redundancy threshold or a performance threshold rather than availability, we may have solved our problem here.

Robert was suggesting though, if we want to keep availability that we need to discuss is having a single K = 1, if that is sufficient for availability and if we don't want to say anything publicly about what are our availability is then we just rename this one and use it with a different name.

DUANE WESSELS: Alright. Thanks for the reminder, Andrew. Yeah, it's time for a break. This was a good discussion and this was using free time. So we can continue having this discussion after the break in the next section. That's what the allotted time is for, so we'll do that. We'll be back here at 10:45.

ANDREW MCCONACHIE: Yeah, 10:45.

DUANE WESSELS: Okay. Thanks.

[BREAK]

DUANE WESSELS: Okay. Welcome back, everyone. This is Duane. Continue to talk about metrics. Before the break we were in the [thick] of this RSS availability proposal. During the break, we spent some time at the whiteboard and there was an enhanced proposal on the table which may address some of the concerns that we've had. What you see on the screen here on the slide on Step 3, it says that for a given probe at a given time, if the measured value of little K is greater than or equal to the big value of K then it was available; otherwise, it was not available. So that is a Boolean. It's either available or not available. And the proposal is to

change that, so that instead of essentially one or zero, you measure at that time interval for that probe a fraction of probes that were available that are less than your threshold, 8.

For example, if that interval 7 were available, 7 out of 8 you use that, whatever works out to be something like 87% is the contribution for that time interval from that probe, and then you add all those up.

UNIDENTIFIED MALE: Does that work in Step 4 mathematically?

DUANE WESSELS: Yeah. Well, Step 4 we'll have to change. Instead of [account], you would have to sum up all those small values.

UNIDENTIFIED MALE: The math behind all of these works even if it's not Boolean.

DUANE WESSELS: So the changes that – in the case [inaudible], you still end up with 100%. But in the case where everything is down, all servers down, you don't end up with zero. Sorry. Still you end up with zero if all servers are down but if you're at the state where you'll have 7 out of 8 for a long period of time, you don't end up with zero. You end up at 89% or something with that. You end up with a value that represents that there is still partial availability of the root server system. But it's below 100%. It's below, yeah.

That's sort of the modified proposal on the table which – having thought about it for a short while seems to work fine for me. We can continue along those lines. Suresh, did you have a comment?

SURESH KRISHNASWAMY: No. It's actually just a clarification whether you're still calling it availability. It looks like we are.

DUANE WESSELS: Yeah. Again, during the break, the discussion that we had at the whiteboard, people who were there said, "If it's this way, I'm fine for it to be called availability." But not everyone is at the whiteboard, so please chime in if you feel differently at this time.

RUSS MUNDY: I don't see Liman back yet. I still have an inclination towards the redundancy threshold or the redundant performance threshold, as opposed to availability because just the term "availability" itself tends to infer essentially a Boolean answer. You either are available or you aren't available. When it's the system as a whole, that doesn't seem to be an accurate description to use the term "availability," but I'm not hard over about it.

DUANE WESSELS: Russ, I can kind of see your point for small time intervals, but the end result of this is a value over the course of a day. And certainly in the RSO case, we're talking about numbers that are not one or zero right there,

there are going to be something in between over the course of a day. Personally, I think it's better if we can have sort of consistency or symmetry between the RSO case and the RSS case, I would feel uncomfortable if we had RSO availability, and RSS something else. But really trying to measure the same thing, I would find that a little weird. Again, I'd welcome other opinions.

I was just responding to Russ's comment about maybe availability is not the right word. I'm sort of arguing that it's better if we can have both RSO availability and RSS availability, and keep the naming consistent.

BRAD VERD: To me, this isn't availability. Availability is a common term that's used in the technology world and it is a percentage. So to say it's Boolean of zero or 100%, to me is odd. I mean, you're kind of getting the best of both worlds in the new equation because your probe or your vantage point is responding in a Boolean fashion, 1 through 8, if 1 isn't there, it's not available. But the other 7 are saying it is available and your accounting for that in the final number, the sum in the end.

RUSS MUNDY: One of the questions I had, Duane, too is that this deals with one vantage point and how the result being a percentage makes it easier than to do math and for the 20 vantage points over the given same five-minute period, is it just a simple math to do the combined –

DUANE WESSELS: Yeah, right. Again, given the little modification to this, here in this sort of terse proposal, we would change the last step to be instead of one count divided by another count, it would be a sum of the case [sub tv] value. So you would just some those up or average them or something and get the value that way.

I apologize that we don't have enough time to put the new … You can run through it on … Do you want to run it through on the whiteboard do you want me to?

BRAD VERD: So here's the way I thought about … You take the 288 intervals, 20 vantage points times the number of 8, which is a whole number, meaning we're saying 8 is 100. Then you end up with … This is not the exact number but it's 46,000 and change. So that equals 100%. Everybody, follow me on that one? But if you do it again, and you do 288 times the 20 probes … I forgot how the math works on this but it's the sum of the 20. So if you had 20 times 7 … I'm sorry. Well, yeah. Let's just do that. So we have 20 probes that responded 7. In the old equation, that'd be zero percent available. Do we all agree with that based upon the old math?

New math, this equals 40,000 answers and then we divide 100. So we divide the 46,000 by 40,000 and we end up with 89%. So in the event of 24 hours a day for 280 probes – this is a 24-hour window – in a 24-hour day where there's a big attack happening and only 7 roots were responding, the equation originally proposed would be unavailable because it was Boolean. In this proposal, it would be 89% availability.

DUANE WESSELS: Yeah. I think when we add this to the documents, we should probably include some examples like this to help people understand how we expect it to work. That would be very helpful to provide some concrete examples like Brad has done on the board here and say, "This is what you get when 7 are available for the day or whatever."

It sounds like we're sort of selling this. The other part of … Oh, Suresh, go ahead.

SURESH KRISHNASWAMY: Just a question. What if we had 9 or 10 that responded, would that number become greater than 100?

DUANE WESSELS: Yeah. In any interval where you have more than 8, you sort of cap it at 8. So it would never get above 100, this metric.

Okay. The other thing we need to discuss for this metric is a threshold value. Now that we have a formula, we should think about what the threshold should be. My guess is that it should be close to 100%. Let's see what spreadsheet says. The answers provided via the spreadsheet were based on a much different calculation of the metrics, so some of those answers are probably not relevant anymore. But looking at it, I do see quite a few that say 100%.

I don't know if people think this should be exactly 100% or if some number of 9s. When we were at the whiteboard, we did a sort of simple

example. I lost it in my calculator but let's see if I can recreate it. 760 times 8. Based on this math, assuming 20 probes, the largest availability you can have that's below 100 is four 9s, essentially 99.99. So that's 1 probe in 1 interval, reaching only 7 servers instead of 8 servers. If that happens, you get 99.997% availability.

UNIDENTIFIED MALE: That's just for a five-minute interval, and then they're going to be averaged over the whole day. Is that correct?

DUANE WESSELS: They are average. But again, what I'm saying is that given this formula, there are no availability values between four nines and 100%. So the first drop down is at 99.99.

UNIDENTIFIED MALE: So just for clarification, when I put 100, it was 100 with K = 1.

DUANE WESSELS: This is 100 with K = 8.

UNIDENTIFIED MALE: When I put 100, I had K = 1.

DUANE WESSELS: In the spreadsheet you're talking about? I see what you're saying. Okay.

UNIDENTIFIED MALE: I think any of the numbers we pick are fine because these are alerts to us. They are not to the outside world. Quite frankly, we can pick it really high, but then I think we have to do a postmortem on every time that we got red. We could pick it low and then have people attack us or, my God, do you really think that the root server system has been perfect over the last year? You didn't have a single one … I mean, I think at this point, we're just gaming how we have to respond to the outside world.

But I will make a significant point here, which is we keep talking about one number where it's four. It's over the four transports and we do see variability over the four transports. Some of that variability, I am very confident even though I have absolutely no data behind it is all about either routing or other transport issues between the vantage point and the server has nothing to do with us.

So I don't know how we will deal with that but I would hope that we would set the number low enough to be believable so that we get one or two of them a year so that we can write reports on them. Again, assuming this is public data, it would be valuable because I think a valuable report is … Oh no, everything was really fine. There was clearly a routing issue here that happened, so this is an artifact of the way we collect by saying availability, and it was really, really available to almost everybody. It wasn't available to that vantage point, particularly if the routing sucked near the vantage point, not near the instance it was going to. But many times, the routing will suck near one instance for a few hours until somebody at wherever it is kicks the router there. I think that those are reasonable things to report.

DUANE WESSELS: But to be clear, one routing problem by one instance is not going to trigger this threshold by any means. It would have to be five. Five servers would have to have routing problems at the same time to –

UNIDENTIFIED MALE: Or five vantage points.

DUANE WESSELS: I think we still have work to do in the document to talk about vantage points when they should toss out their measurement, when there's a problem close to a vantage point and when those measurements should be discarded. That is not as well defined at the document as I would like so we have some work to do there. But the intention is that if it's a vantage point problem, yeah, those measurements should be thrown out.

PAUL: I'm sorry, this is Paul again. But to be clear, from at least the anecdotal data, which is all I have, I believe that might be a hard determination between TCP and UDP. It seems like for some of them, UDP goes to hell and TCP is fine. You would expect UDP might be a little bit bad and TCP was fine but for some of the stuff I've seen, for some reason on a day, UDP goes to hell and TCP is fine. That doesn't seem as likely to me. Again, I think we may have to write postmortems that feel like hand-waving, but in fact, are explaining measurement.

RUSS MUNDY: The document says the public reporting is on a daily basis. So that's going to really camp down problems as far as what gets reported publicly. But again, we may want to have more internally visible things like … but that's something that can be addressed, I think, in the future. But with the threshold of 8 and this number of probes, probably it seems like for our threshold target, we ought to be looking at if it's 99.99. That's probably what we should use for our public stated threshold it seems like.

UNIDENTIFIED MALE: Complete side tangent, but it just came to me. Publishing, it seems like the assumption for the parts of these metrics are going to be used by RSSAC 038, it's the PMMF++, will be doing the publishing. Who's going to publish the RSS metrics? Do we want them doing that? Do we want us? We don't have a publishing mechanism within RSSAC other than … I don't think we want to burn RSSAC 0 whatever number for every day. I mean, are we thinking that the RSS metrics and thresholds that we're talking about now are going to be published by this PMMF++ or is that going to come out of here? Nothing technical, but going back to Russ's question, internal publishing versus external, we should think about that at some point.

DUANE WESSELS: I know. I think that's still TBD. But one thing this work party can think about or decide is that there was a proposal yesterday in the case of the

RSS metrics to publish the actual values rather than a pass-fail indication. So that's something that we can continue to think about.

BRAD VERD: Something to consider. I'm not passionate about not publishing or publishing, but I feel that the first request we get for the data, if we're not publishing, we're going to have to have a good story not to share it and I don't know what that story is.

DUANE WESSELS: Do you mean the raw data or just …? Okay. Well, one reason that we can talk about now, I guess, good as time as any is, the work party settled on this idea of publishing pass-fail only to discourage gaming of the system. We felt that – especially in the case of latency, if the actual latency numbers were to be published, that might encourage operators to do silly things like put their servers where known probes were so they could look better in terms of latency. That was the main reason for that.

BRAD VERD: If gaming the system needs more instances, I'm okay with that.

UNIDENTIFIED MALE: That might not necessarily be the case. If you just have your 20 instances in the right place, why would you need 25? If they all close the probes and everything looks really good, that could be –

BRAD VERD: The old statement. Why would you?

DUANE WESSELS: All right. Should we get back on topic to the discussion about the actual threshold for this RSS availability metric? Unfortunately, as we've talked about, this spreadsheet was kind of filled out, given them a different proposal, so some of these numbers are not good. Would anyone like to throw out some suggestions for what they think this threshold should be set up?

Sure. The highest example that you can get that's not 100% is 99.99. It is possible, as we discussed at the whiteboard, if all of our worst fears come true and there's an attack that lasts longer than a day and somehow manages to take out every root server everywhere, you would get zero. So zero is a possible output of this formula. Very, very unlikely obviously. If you lost, say, half of the root servers, you would get something like what's up on the board. Well, it says 89%. Let me do the math. So if you were in a state where for a whole day, you had only seven root server operators functioning, you would get 87%.

LARS-JOHAN LIMAN: I'm more curious about when you lose much for a fraction of a day instead of cutting it this way, if you cut it the other direction, how does that impact the numbers?

DUANE WESSELS: That didn't work. So here's an example, if there was an attack that's less than five minutes, so it fell within one- and five-minute in interval and took out all the root servers, you would get 99.6%.

LARS-JOHAN LIMAN: So with that, one extreme example, I would say 99% seems like a reasonable threshold for us to be alerted. Again, this is an alert threshold. This is not a massive failure threshold. We need to look at it, we need to look at the data and figure out why. Quite frankly, I think we're going to need to in that example, because the RSO numbers are going to go to hell around there as well.

I think we should be alerted – we, as RSSAC – when something like that happens or there's a periodic one during the day, even if someone's only taking the entire root server system for a minute, just to see what happens then comes back, that will probably still fall into that 99% threshold.

RUSS MUNDY: 99 as the public published threshold for the RSS seems very reasonable because it does permit some problems to occur that might or might not show up as red and would show some variation in the numbers over time. But it would, as you say in your examples, it's still take a quite sizable successful type of attack on either the vantage points or the root servers themselves. But 99 is as far as a target threshold number for what we believe the RSS should be available. I find that there will always be people that can come up with a way to complain, but 99% is pretty hard to complain about.

UNIDENTIFIED MALE: So 99% for a day, fine. But if we have 10% for five minutes, in Liman's case, we should definitely cause for … maybe reportable a day but for our own purposes, looking at five-minute intervals, things that are drastically low is important.

LARS-JOHAN LIMAN: Yes, I agree. Exactly that. What we're talking about here is designing a system where we fail to comply in a contract with the public, so to speak, and that failure should lead to a public report about why something happened. The numbers are still there, they're still available and they could generate any number of actions just by looking at them and interpreting them in different ways by – I would argue with you, you would say by anyone – preferably by the root server operators and RSSAC in the first hand, but the numbers should be there, it should be able to look at them and do calculations and come up with actions that may need to happen but may not necessarily generate a fault as doing report in this context here.

DUANE WESSELS: Okay. What you just mentioned sounds a little bit to me, like monitoring rather than a metric thing. You said you wanted to be notified if there's a problem in some five-minute interval. Were you thinking like notified right away or just notified later?

UNIDENTIFIED MALE: I don't know but it's certainly something for concern, because it take up all RSOs for 5.1 minutes. Our overall metrics can still look good, 99%. But that is a problem.

DUANE WESSELS: That can be addressed with a different threshold over a different period of time. I just want to make sure that what you're not talking about is real time monitoring, because the work party document doesn't say anything about that at all at this point.

UNIDENTIFIED MALE: Right. A report may need to be generated at some point. We will have those numbers available to us for smaller time intervals and they'd be actionable, maybe outside of the metrics.

RUSS MUNDY: Ken, do you think this would fall under the purview of 4.8 the unexpected results to have an event of this nature? Because we do have words in there that says the information needs to be held on to for later analysis by other –

[KENNETH RENARD]: I think yes, definitely.

RUSS MUNDY: Because having everything down for any length of time has happened. So I think it would tend to fall in that category but it's more a design of

the implementation and who's actually going to be doing it and the internal data availability, I think you're going to see. Is that right?

[KENNETH RENARD]: Yeah.

UNIDENTIFIED MALE: So this brings up another meta question, which we don't have to fully do here, but the way that the document is currently structured, it seems – and again, this was based partially on my misunderstanding of how much of this is for 038 and how much of this is for us. But at this point, especially because the no RSO should be seeing each other's data, I get the feeling from this document and the way I've been working is that whoever the central collector is has all of the data, and that they would piecemeal it out if someone said, "Hey, I failed. Why did I fail?" What you're asking for, I think, here is something where RSSAC would have all of the data.

Now, again, I don't really worry about that. As an – ooh, you're going to know what your competitors are doing because you can know it anyways with any other probing system. But there might be something that we want to put in there is that whoever is doing the collecting, once it comes off the vantage points and goes to a central place, also must share it with RSSAC in a timely fashion if we are going to be concerned about things like that. So that there would be a second repository that RSSAC could be running its own evaluation over. I certainly wouldn't want us to have to ask whoever was the PMMF++ to also "Can you run this program over the data so we can see this up?" I

think it would be better if they up to date copy with sitting with us. Does that sound reasonable?

DUANE WESSELS: Well, I think that – geez. I'm sorry.

UNIDENTIFIED MALE: Don't touch it.

DUANE WESSELS: I think it's important thing to discuss, because, again, based on discussions in the work party a few months ago, I sensed a lot of sensitivity about making comparisons to each other and gaming the system. If RSSAC has a copy of the raw data, then of course, the root server operators can compare themselves to each other.

I don't really see a way that we could … Since the RSS data is built on the RSO data, you can't separate those two. You can't sort of share one and not the other. So I guess we would need to hear from people whether they want this ability to see the raw data or not.

BRAD VERD: Can the people who are concerned speak up?

DUANE WESSELS: Let me ask that … Sorry, I didn't see.

UNIDENTIFIED MALE: I should sit closer. The data we're talking about does not have PII in the usual GDPR. Since we're not collecting end user IP information, we may be talking to a resolver that has a very small population that's making it possible to take our data, and deanonymize it down to the human who made a query. But we're doing the best we can not to have that be the case. Certainly, queries coming to us from Google could have come from almost any person on earth. Also, our intent is not commercial or otherwise controversial at all.

I live in DNS collection world and I can tell you that what we are planning to do is so much less controversial than what Google and Cloudflare and Quad9 are already doing, then I don't think anybody is going to raise an eyebrow. Thank you.

BRAD VERD: I think the people who are concerned need to speak up and/or we're not going to just make the data available fully transparent. I think that there is a justification that needs to be written so that when there is an outside party who asked for that data, it can be given. Because I stated earlier, I don't know what that story is. Fred, that'll be somebody asking you for that data, you need to know what to say back to them when the answer is no and they're going to say, "Why?" So I think it's easy to say, "Data is available."

RUSS MUNDY: One of the things that – we're sort of going, I think back and forth between some of the exceptional events that Kevin was talking about earlier, the data that gets collected on an ongoing basis. So we're not totally in the 4.8 realm, but in the 4.8 realm, for the bad things, the

unexpected events, there we have some words, there should be some additional things perhaps collected made available for subsequent analysis. But the data that at least it's my perception that we've defined that we're collecting is from a sensitivity basis, very innocuous sort of data. I'm, at this moment, in Fred's camp for all the routine data. I would urge that we just say it is available if people want it and they want to do their own analysis of it and so forth.

BRAD VERD: Just to be clear, to restate what I said, I'm not passionate either way. I'm not advocating for it to be public, I'm not advocating for it to be private. I'm just saying that sitting in the seat, and now Fred's going to be in this seat going forward, is we will be asked and so the story needs to be written. If we're going to say no, why? Right now, I can't answer that. So in lieu of not answering it, the easy way to address it is to say yes.

UNIDENTIFIED MALE: I would argue that the right implementation of that would be to simply make it public.

UNIDENTIFIED FEMALE: I don't know if this is helpful or not, but in a parallel context where we track data for the CSC and public on our performance on the root zone management. In the dashboard, we do make all this data available. When we display it, it's anonymized for each of the tests that were done and how the performance was. However, we do have an option to

download the role log of the data that we used to make all these calculations.

We do anonymize the data so that you can't trace it back to us in TLD and see that they consistently failed X or Y tests. But if the data is available, because that was a mandated thing that we want to check how ICANN is checking this and publishing the data. So that's something that we've gone through and we've made that available.

People, they don't even have to request it, it's available in the dashboard. Then I think the last thing I wanted to say here is I'm in support of Brad. I think the data needs to be out there and let people analyze as they wish.

DUANE WESSELS: So definitely, we need to add some text to the document about this publication and data sharing. Again, I feel like right now, we're leaning towards being very open and transparent about sharing data. But in the past, we've felt the other way. So I don't know yet how we're going to settle on the final thing that goes in the document but I guess we'll figure that out. I mean, one thing we can do is we can just make proposal to be transparent and throw it up and let people object to it if they want to. Maybe we'll take that approach.

BRAD VERD: Again, if they object, they need to provide the story so Fred's got something to share.

DUANE WESSELS: So lunch is in 45 minutes and we're kind of way ahead of schedule right now, I think. The next thing on our –

UNIDENTIFIED MALE: No. We have latency in this section.

DUANE WESSELS: Yes, I guess you're right. That was not in my notes so I think I missed that. Let me go to the document. Okay. So moving on from RSS availability, the next section would be RSS response latency. So this is Section 6.2. We talked about this a little bit yesterday. In the document as written … Again, all the individual root server operator measurements are just aggregated into a value for the RSS as a whole. As written, it's a very, very simplistic aggregation method, which is to lump all of the measurements together and take their median value.

We had a proposal from Shinta yesterday to not do this. Is Shinta online again today? If I remember correctly, one of his proposals was to consider instead of 50th percentile, maybe 30th percentile or something like that. I have a lot of sympathy for this line of reasoning. To me, it's not particularly useful if we sort of say that the latency of the system as a whole is simply the average or the median of the latency by the individual operators. That's sort of too simple and doesn't really provide any additional useful information to the metrics. Kevin has got a –

RYAN STEPHENSON: Sorry about that. I just wanted to look at that and see what RSSAC037 says about any reports that are developed by the PMMF. There actually is a paragraph in there in addition to measurements of individual operators, the PMMF develops and implements techniques for monitoring the RSS, such monitoring may lead to reports. Example, IANA reporting on performance which, of course, Naela brought up, on average response times and availability for the system as a whole. So to just go in line with RSSAC037, by making these measurements available, it keeps it in line with that document. Just FYI.

DUANE WESSELS: Thanks, Ryan. I think that the text that you just read in in that document is sort of general enough that we could satisfy that whether or not the raw data gets published. I think that's sort of the open question is certainly there will be reports. What we're proposing for the reports is I think the tech specifically referenced latency numbers, were proposing pass-fail on latency threshold was a little bit different. But the real question, I think that the group needs to think about it is for the raw data, is that to be made openly available or not.

UNIDENTIFIED MALE: So, Ken, you had a comment about the response latency?

[KENNETH RENARD]: Yes, I can. I agree that the average of the medians is maybe a little too simplistic. Just throwing another idea on the table along with the 30% or 30th percentile threshold. Trying to emulate resolver behavior, if

they're going to prefer closer or better ones, do we take the top end where we have a nice value of K = 8? Maybe we reduce to the best 8 medians and use those? Or some variation of that thrown out there for [discussion].

DUANE WESSELS: I think that's very tempting. What makes me a little bit uncomfortable is that we've kind of taken steps to a document to say that we're not really about resolver behaviors and that sort of thing. So that's where we have the challenge there, I think. It's almost unavoidable though because if you are totally ignorant of resolver behavior, then about the best we can do is this median of the whole bunch thing. So I think we have to find the right balance. We have to venture into that territory a little bit but I wouldn't want us to go so far as to say, "We know, resolver works this way and it favors blah, blah, blah." I think that would be too detailed and too much.

UNIDENTIFIED MALE: Going down the path of resolver behaviors however, since we do know that many resolvers do their picking by buckets. So it doesn't have to be the top aid, it could just be … So this would not be a percentage RSS response latency that I'm proposing. I've thought about it for all of 30 seconds, but percentage of RSOs whose average latency was below X, because that is the way that many resolvers deal with it. That is not like who is the lowest but who's below this. We could pick a number, like 15 milliseconds or 100 milliseconds. That's clearly not an average and it'd be very explicitly not an average, but if this is number for us then I think

the desire here is to have RSOs at least be responsive to one population on some shared network. Responsive means you might get picked because you're fast enough. Again, I think we could just pick a number and look at the percentage of the RSOs that day who were better than that number.

DUANE WESSELS: Okay. You are proposing something which would be a big change of this metric, instead of the result being a latency in terms of milliseconds, it's percentage of servers who's likely to meet some criterion. To match those, we come up with a threshold so we would have to come up with a threshold that says, "Okay, 10 RSOs must have latency below…" or "8 must have latency below X to be in a…" Well, there's that word good again.

UNIDENTIFIED MALE: That's what the RSO specials are all about is good.

DUANE WESSELS: I'm willing to consider it.

RUSS MUNDY: I agree, we certainly could do this. One of the things that struck me about the suggestion though, is if we describe latency in terms of a percentage for people looking at it, that's not a logical measure that you apply to latency. If we do go this way, I think we better have a really good explanation in the document and be prepared to provide more as

to why we went that way. I mean, people, when they see latency, they think an absolute number.

DUANE WESSELS: Let's give Shinta a chance to talk if he's on the Zoom channel.

SHINTA SATO: Yes, I'm here. As I said in the comment in the documents, I did not think that the median represents the number of the response latency. However, I'm not sure what would be the best way to determine what the metric should be, so I just wrote that the minimum number would be the best for the one case. Maybe 10 percentile, then 30 percentile or something like that would be good but I'm not sure what it should be right now. I'm sorry about that.

DUANE WESSELS: All right. Thank you, Shinta.

SHINTA SATO: Yes, we may need more discussion here.

DUANE WESSELS: Okay. Liman has a hand up.

LARS-JOHAN LIMAN: Liman here. If we get into these numbers, what we're doing, we are creating complex metrics that consists of aggregated metrics from

measurement and integrate it over time. I don't really have a gut feeling for what these all mean. So I think we could do less of haggling here now and say that we go with one prototype. I think we have to set up the test probes on the real Internet to see what the numbers are to get a feeling for what are we measuring, what can we measure, what happens when we combine them, what happened when integrate them over time, what actual numbers pop out to the other end. It could be that we could do simulations to do this but I can't really bend my head around these things and say, "That was not open. 09998, that'll be the right number." Because I don't have a gut feeling for what the numbers are for real input.

DUANE WESSELS: So I would certainly argue that we set a lot of thresholds for RSO latency, 250, 500. If we're going to have the exact same threshold for RSS latency, then the current definition is perfectly adequate. There's no reason to think that the median of the group threshold would be any different than the median of an individual RSO. So the only way that it would make sense to me to consider a different measurement technique for RSS latency is if we were willing to have a different threshold, a lower threshold and I'm not sure that we're willing to do that. So maybe that's a question to put to the group. Would people be willing to stick their necks out a little bit and say, "RSS latencies should be on average better than individual RSO latency."

LARS-JOHAN LIMAN: I think no because if you stick your neck out to say that, what you're actually saying is that I'm willing to take the responsibility to make sure that this is lower by providing a better service. If I'm willing to do that, then we should probably lower the number of RSOs ourselves as well. So this comes back to another number where you said that we needed the same number for the RSS system. If it is a minimum performance, then that is the minimum performance for the operator and for the system as a whole. I fail to see difference there in this case. There could be other cases where you could argue that you can take a chance that the numbers will play in our favor, which doesn't seem like the right thing to do from my perspective.

DUANE WESSELS: I guess I just want to quickly throw out, based on where we're settling in particular about the RSO latency thresholds, where we settle there, and looking at real data that we get from Paul's prototype [inaudible], there's no way we're ever going to exceed these threshold. These are very generous thresholds. Unlike some of the other things like [correctness], I think, yeah, there may be a possibility that we could be below our threshold there. To me, these thresholds are so high that we're almost never going to exceed them.

UNIDENTIFIED MALE: So why even have an RSS latency?

DUANE WESSELS: That's that that's a good question. Again, if they're going to be the same, if RSS latency is the same as RSO latency, then maybe we don't need both. I don't know.

Suresh, go ahead.

SURESH KRISHNASWAMY: So when we discussed the RSS availability metric, we sort of agreed that there is a reportable metric that the world sees and there's one more useful for the RSO community. Similarly, I think there is one for the RSS, because you want to report something that is I think the median is fine as an externally reportable value for the system. But in terms of what will be useful for the root server operator community, is to know that there is a systemic problem that is affecting latency for more than one operator. The right way, I think, to look at that, going back to a comment I made yesterday is to look at the spread and something like a standard deviation on how much of disparity do you see in the latency is across all of the servers at a given point in time? So I think that measure is probably more useful, has an internal reporting thing but I think it might be getting there.

DUANE WESSELS: Thanks, Sirush. At a previous meeting like this, I presented some graphs and some data that had convinced me that in these cases where we have distributions, where we'd look at medians and these long-tailed distributions, standard deviation itself isn't a good measure of spread, a better one is just the percentile range. If that's the case, then maybe as Shinta is suggesting, we need not at 50th percentile but a lower

percentile as our threshold. That's where I would take that conclusion, I think,

LARS-JOHAN LIMAN: What is the lower percentile? In which direction?

DUANE WESSELS: Well, the stuff that I presented last meeting, the point was, instead of looking at standard deviation as a measure of precision, you look at a percentile range. Maybe 25th percentile to 75th percentile.

LARS-JOHAN LIMAN: That's stopping where? At zero milliseconds or at infinity?

DUANE WESSELS: If you take all your individual measurements, lump them together, and you calculate the 25th percentile.

LARS-JOHAN LIMAN: 25th percentile towards zero. Or the 30th is a tighter thing than 50.

DUANE WESSELS: Yes. 30 percentile is more aggressive or more stringent requirement than 25. You could also look at the other way, look at the 70th percentile to say that we don't want numbers that are too high. We don't want a lot of numbers that are too high. So I guess that's where I would take that in the direction of looking at the spread.

BRAD VERD: I think the question was really, I heard somebody say, "Maybe we don't need to have a threshold for the latency of the root server system." What I said yesterday, I still kind of believe it is that … I said this during the latency discussion, which was when we agreed on – was it 250, 500, super high. That maybe we add some documentation in there or some verbiage that says our goal would be that if an internet user is beyond 150 milliseconds to the service, meaning any route that it can reach an individual server, that they should reach out to us or something like that. But trying to measure that from the user basis, not reasonable.

UNIDENTIFIED MALE: I'm not going to touch it, I'm just going to lean over. Worked this time. So you had said that yesterday when we were talking about RSO, did you mean maybe to put that text in the RSS section? Great. Because I think that that makes more sense in the RSS section.

At which point, either we might have no threshold, or we say, "Here's just a straight average." Which we know isn't a good measurement here but it's useful as a way of having parallelism between RSO and RSS and say, "We'll just do the averages. We don't know what that means." We won't necessarily have a threshold there until we've done more investigation and such, but we can just say, "Here's the average." and there isn't a threshold for us at this point. I'm sorry. Plus wording.

LARS-JOHAN LIMAN: If we don't know what it means, if we don't have a threshold, it has no room in this context. We can still do it, we can play with the numbers but if we don't use it for measuring minimum performance, then it doesn't have a place in this context, I think.

UNIDENTIFIED MALE: Then we just say that the minimum … Again, same thresholds as the average for the RSOs so that the threshold would be to 250, 500. Same way.

UNIDENTIFIED MALE: So I think we should have a threshold and I think it should be lower than the RSOs. We're not defining good here, I think what we're defining for the RSO is good enough, which allows for a diversity. I'm not saying that the bigger letters should turn off half their instances and settle back to 250 is good enough. But for the average system as a whole, we want there to be better performance than the lowest performing RSOs. I think there should be a threshold and it should be lower than what the RSOs are.

LARS-JOHAN LIMAN: Following onto that, if we do have that situation, we have a rather lax threshold for the RSOs and a stricter one for the RSS. The root of a system fails to meet its threshold. Where's the burden to fix that?

UNIDENTIFIED MALE: [inaudible]

LARS-JOHAN LIMAN: Okay. Fair enough.

UNIDENTIFIED MALE: I'm supporting what Robert was saying about the argument for different numbers and I don't think they need to be wildly different. We can do better from a location of a vantage point. We think the RSS can do better than a single RSO. You're just going to raising that bar a little bit because the topological geographic diversity of a system versus single or so.

SURESH KRISHNASWAMY: I like that general idea but … because we're trying to make use of the least number of queries sent from the vantage point to the servers and do computations from that. So are we then proposing to establish, for purposes of the report, simply a lower number but determines that number based on the average of all of the response for the RSO that are used for the RSO measurement latencies or are we proposing doing additional probes to random RSOs for the RSS measure?

DUANE WESSELS: There's no proposal for additional measurements.

BRAD VERD: Can you use the existing measurements? I mean, of the 288 that are going to all the different roots, you could take the best one for that five-

minute period. So you get 288 best measurements for every five average data what you're calculating as – whatever your threshold is, that your tool that you're using, are you above or below it?

DUANE WESSELS: So that sounds reasonable. So in a five-minute interval, you take the best K and take the median of the best K over time, and that's your RSS latency. That works. For me?

UNIDENTIFIED MALE: Sure.

DUANE WESSELS: Is that a good starting point for other people? Okay. So thanks for that. That'll be a big change to that metric symbol, we rewrite that section and I don't know if we're prepared to talk about thresholds today. But maybe. Russ?

RUSS MUNDY: Since we really don't know what they would be based on real data but the suggestion is a little bit lower. Should we just pick an arbitrarily amount smaller instead of 150 –

UNIDENTIFIED MALE: 250 and –

UNIDENTIFIED MALE: 150, 300.

RUSS MUNDY: Yeah. Or 200 and 400 or something like that.

UNIDENTIFIED MALE: We hear 150. Do we have another bid?

DUANE WESSELS: Sure. I think, again, there'll be work to rewrite that. Maybe while we do that, we'll throw a number. We'll think of 150 at this point and we'll throw some numbers in the document and then come back to the group and say, "This is what thresholds we're proposing at this time, but open for discussion still if people want to have discussion." In the meantime, we can also ask the people who are doing prototype data collection. So look at the data and see what that number would look like today as a safe comparison. I don't want to do too much of that, I don't want to design too much to the current system but it's good to have a sanity check.

UNIDENTIFIED MALE: Although, as I think the only person who's doing this now, this is all just the report. John, are you working in this realm as well? Great. Then we absolutely do need to talk next week. But this is all just the report at the end. Now that all the data is here on the central one, which things do I collect and put out? That's an hour per thing changed. It's really not a big deal once you have all the data sitting in front of you. So if the work

party said, “Well, what would it be at this thing?” It really would be easy. Quite frankly, as I mentioned to John earlier, I’m planning on making my code available to everyone here. Not that we had much success with that in the resolver work party of anyone doing anything with it but if folks here want to take the reports on doing and change a thing, that would be just fine as well. So I’m not hesitant on the, “We’d have to go back to John and Paul on this.” That’s a fairly easy thing.

SURESH KRISHNASWAMY: We sort of went back on the idea. The idea was proposed that do we need any additional measurements to test the fact that the RSS measure is actually lower in practice. Then we decided we’re not doing anything new. I think the fact of the matter is that, in reality, a lot of how latency will be measured is going to depend on resolver behavior. Assuming that it’s going to pick the top … So the measure that we using for the RSS latency is just based on the best K. So that satisfies our requirement that it’d be better than the RSO metric. I don’t think we actually getting to, “Is it really better in practice?” I’m asking the question, again, “Do we want to reconsider?” Whether we want to add that test where a vantage point sensor a probe is going to be … in order to resolve a query, is going to have the opportunity to look at multiple resource to resolve, and then use that maybe just internally in order to see if there’s a better measure for the RSS latency but just have that measure available.

UNIDENTIFIED MALE: Did you understand what he was asking? Because I think I did and it's cool. I think to say that quickly, is that instead of looking at the top eight in a day for each five-minute period from each vantage point, what if that vantage point was a resolver, which would have chosen and that's – well, it's not binary, it's whatever you would call binary 413. Pick the root server, it would have done and just report that. That's cool.

SURESH KRISHNASWAMY: That's not exactly what I suggested but I like that better.

UNIDENTIFIED MALE: We can do that. Nothing in what we're doing now says we have to come out with just one number. As we talked about earlier for availability, we might have publicly available number and then internally looking, for example, for an hour and such like that. We can do these in addition. Once we have a data set in front of us, we can generate our own internal reports based on things like that, as well.

DUANE WESSELS: My concern with that approach is that you have to very accurately model the behavior of all the different resolvers or you have to make a lot of simplifying assumptions. We think we know how they work or how they should work but we're not sure how they really work. The service selection changes in a given implementation that changes over time, it may work this way and when the software is released in 1998, that works this way. I think there's a lot of complexity there. But if

you're willing to make certain [simplifying] assumptions, then you could probably do something like that.

UNIDENTIFIED MALE: Duane has a data point for what you just said. Google public DNS, which is the one that if we're counting users is the most important, changes their attraction model over time. Actually, some of their nodes on a given day are using two different traction models. So we really, really – you can say much more definitively, we don't know how they act, but with some simplifying assumptions, saying, "If they acted like this, this would be an average latency. If they act like this…" So we might have multiple answers. That doesn't come up with what one do we want to report to the public but we could have many ones for internal as a way of saying, "How are we doing as a root server system?"

DUANE WESSELS: I guess the question I would have is, "Do we have the appetite to do this in this version of the document or is this future work?"

UNIDENTIFIED MALE: I'm assuming it's all future work but it could be ongoing future.

DUANE WESSELS: I don't know if that matches your expectation, Suresh. But to me, this seems like a lot of complexity to consider.

SURESH KRISHNASWAMY: But I think it's more useful in terms of a metric but I'm fine with having some [later].

DUANE WESSELS: So what we should at least do is … I mean, again, as always, document our discussion here and make a recommendation for future work along these lines that we should consider doing a better job of modeling resolver behavior in a future version of this –

UNIDENTIFIED MALE: Or even if not a modeling behavior. At least give us ourselves more internal metrics for us to determine how do we feel that the roots are … I mean, this is all about good for the root server system. So we might change our view of good over time. What a five-year-old considers good for family life is very different than what a 13-year-old does, right?

DUANE WESSELS: Okay. Anyone else that have comments or questions about RSS latency? Is the food here?

BRAD VERD: Yeah. Food's here, food's ready. There's three catering trays out there. First one is chicken, second one is potatoes, third one is beans and what looks like chicken but it's vegetarian chicken. I don't know what that means but those of you who are looking vegetarian, that's your go. Okay?

DUANE WESSELS: All right. I think this is a good time for us to take a break. We'll take a lunch break and then by the schedule, we're back here at 1:30. So an hour and a half. Alright?

UNIDENTIFIED MALE: Thanks, everyone.

**[END OF TRANSCRIPTION]**