DANIELLE RUTHERFORD: Good morning, good afternoon, and good evening, everyone. This is the RSS metrics Work Party teleconference held on the 29th of October at 17:00 UTC. On the call today, we have Duane Wessels, Fred Baker, Jack Biesiadecki, Jeffrey Osborn, Kazunori Fujiwara, Keith Bluestein, Ken Renard, Paul Hoffman, Ray Bellis, Russ Mundy, Shinta Sato, and Kevin Wright. I see an unidentified 614 number, US country code.

From staff, we have Steve Sheng, Andrew McConachie, and myself, Danielle Rutherford. A friendly reminder for everyone to please state your names before speaking for transcription purposes. Thanks. Back over to you, Duane.

DUANE WESSELS: All right. Thank you. Yeah. The plan for today's call is to go through the document and look at the pending comments and outstanding actions. There were a lot of comments added, a lot of suggestions added, in the last couple of weeks. Yesterday, Russ and I spent some time with the ICANN staff to go through and tidy up the document. We accepted the changes that were obvious and easy to accept. Some of the other issues, we left for discussion on today's call.

I think everyone should be able to see the screen that I'm sharing. I'll scroll through the document and we'll talk about the issues. I'm a little bit concerned about time so if you would like to make a comment, please raise your hand or jump in but please try to be brief so we can get through all of the things we need to get through.

Note: The following is the output resulting from transcribing an audio file into a word/text document. Although the transcription is largely accurate, in some cases may be incomplete or inaccurate due to inaudible passages and grammatical corrections. It is posted as an aid to the original audio file, but should not be treated as an authoritative record.

Here, first, we're in section two, the background to scope. There are some comments from Daniel Migault. Is Daniel on the call? I didn't remember. It doesn't look like it. He won't be able to speak to this but Daniel and I are having a conversation in the comments, here. Since he's not here, I'm not sure what there is to say, additionally. Unless anyone has happened to read his comments and would like to make a comment at this point? All right. I think we'll move on. Paul, your hand is up, I see?

PAUL HOFFMAN: Yeah. In the case of these long comments, I'm quite hesitant to add them in. I don't think they actually help. I think they restate some of the things that we have stated in other parts of the document and, in some cases, I think they make them a little bit muddier. I wouldn't want to add this large section unless we understood what the deficiencies in the current document are. And if so, it might be that we can just add a sentence or two where there's a current deficiency.

DUANE WESSELS: Yeah. I agree. My comments to Daniel were that this stuff belongs elsewhere, if at all. And I agree that it doesn't really need to be this long. All right. Let's continue down through the document. Here, we're in section three, which is about vanish points. If I remember correctly, Daniel had added some text, here, to the first paragraph around the word that's highlighted, "manageability." I rejected his proposed text but tried to capture the same sentiment in the new paragraph.

There's a new paragraph here which essentially says that there could be a concern that, as the number of vanish points increases, you may have to deal with homogeneity issues. That can complicate the interpretation of metrics. But it ends with, “As we gain experience in the operation of these metrics, a future update may want to recommend a larger number of vanish points than 20.” Daniel seems to be okay with that. Any comments with this paragraph, at this time?

RAY BELLIS: Yeah. What does he mean by homogeneity, in this context? I'm actually concerned that 20 is far too few and that it could cause more errors rather than see a smoother distribution.

DUANE WESSELS: Homogeneity was my word. Daniel’s words were along the lines of that, if you have more vantage points, interpretation can become trickier. The kind of thing that I was thinking of with homogeneity, Ray, was, for example, in the RIPE Atlas system you have probes and anchors. As you’re looking at results, you may wonder, “Well why, when I look at probes, do I see one thing and when I see anchors I see a different thing?” In my mind, the extent to which all of the measurement systems can be the same, the results are easier to interpret. That includes things like having them all located in similar types of data centers and things like that.

I do agree with you that, generally, more is better, due to the law of large numbers thing. But at this point, I feel like we settled on 20 as our recommendation.

RAY BELLIS: Indeed. It’s just not clear to me what that paragraph even means, as phrased. The concept of homogeneity and heterogeneity, as written in the paragraph, are quite nebulous and don’t really correspond with what you just said, to my mind.

DUANE WESSELS: Yeah. Right. I was reluctant to mention specific things like RIPE Atlas. I don't know. We can propose to strike it or we can make it longer and more verbose. Paul, you have a comment?

PAUL HOFFMAN: Yeah. I would be okay with striking this. I did not have the problem that Daniel had that started this, which was, at least from his first message, that there might be a homogeneity problem or that we weren’t dealing with it strongly enough. I think that by saying “start with 20 and we might be adding more later,” or whatever, that’s sufficient for getting this going. We don’t have to explain all of the possible problems with various ways of doing it. I think the more that we go into that, the less clear the number. Again, this section is on the number of vantage points. I think we can, basically, just say, “20,” without saying, “and we chose that number carefully,” because we didn’t.

DUANE WESSELS: Okay, thanks. Russ?

RUSS MUNDY: Yeah. Just taking a quick look at what's here, I think we can take the first existing four-line paragraph and, maybe, add the last sentence of the paragraph you just added. It would then end with, "As experience is gained in operation and interpretation metrics, a future update may recommend a larger number of vantage points."

RAY BELLIS: That's exactly what I was going to propose, as well.

DUANE WESSELS: Okay. So, at least for the people on the call here, the consensus would be to strike that part and keep the last sentence, right?

RUSS MUNDY: Yeah.

DUANE WESSELS: Okay. All right. Thank you. Let's move on, then. We have a discussion here about spoofing. The comments are very long here, so they span the whole heights of the window. Section 4.5 is about spoofing protections. There's a comment from Kazunori about that maybe there should be new text that says, "Measurements or vantage points should be disabled if they detect spoofing activity." Kazunori, do you want to speak to this comment? I see you're on the call.

KAZUNORI FUJIWARA: Hello. I did [inaudible] could recognize [inaudible] if the root server is responding with the incorrect data or [illegal participation].

DUANE WESSELS: All right. Thank you. So you're proposing an addition to section 5.3, which is specific to the correctness metric. I guess the question which we should maybe take offline is the extent to which a vanishing point should try to detect the spoofing or how it would actually do that. Paul, you have a comment?

PAUL HOFFMAN: Yeah. It was exactly what you just said, which is that at this point we, as a group, had decided that the vantage points are not doing their own analysis and that the analysis is going to be done by the collector system. But further than that, I have a deeper problem with what Kazunori is suggesting, which is, let's say that we even said, "Oh, well, if the collector system detects spoofing then the vantage point should be taken out. And if it detects spoofing by correctness …"

But that's exactly what the correctness checking is supposed to do. A vantage point, or the main collector, cannot detect whether a wrong answer is spoofing or a root server issuing a wrong answer. I think that we would become a little bit insane if we tried to differentiate between those two.

DUANE WESSELS: Yeah. I mean, it kind of depends on how you write your software. There are some things that you could detect. For example, if you did the casemixing of query names, you could detect and report a response where the case didn't match. You would receive that response and say, "Oh, the case didn't match." But other sorts of things like spoofing port numbers, your application wouldn't even necessarily receive a response or a message with the wrong port number. You wouldn't see it at all. Some things are easier to see than others, depending on how it's written.

I added a paragraph here, at the end of 4.5, which I thought was a way to compromise around this, which has said that, "If and when a vantage point detects spoofing, it should be recorded and logged so that, if you go and later manually inspect those responses, you can disregard them if necessary." I'm a little bit nervous about saying that vantage points should just be disabled outright because then an attacker could, probably, intentionally disable a lot of vantage points, if they wanted to. They could just take the whole system out. Any last comments about this before we move on? All right.

Now, we're in 4.8, here. Paul, this was a to-do item, right, which we haven't gotten to yet?

PAUL HOFFMAN: That's correct. I didn't actually get it assigned to me until yesterday during the infamous California power outage.

DUANE WESSELS: Okay. All right. This is work that still needs to be done. Or, alternatively, we can be less specific here and say, "We expect the builders of this system to handle this appropriately," or we can be very descriptive and say how we expect them to handle vanish points that don't seem to be functioning well.

PAUL HOFFMAN: I'll take a stab at this in the next two days and am happy to then have folks [beat] on that when we meet a few days from now.

DUANE WESSELS: Okay. Thanks. We're now in the section which talks about how we came up with this formula for determining the number of RSOs required for reliable operation. This is the K-out-of-N model, and this is how we came up with our value of K. There was a suggested addition from Shinta, which is talking about how the root server operators have different deployment strategies. I tried to take his sentiment and rewrite it a little bit in the green paragraph, here.

One thing I didn't quite get was the last part of his paragraph, which talks about how the current measurements do not describe the difference when K RSOs are available. Shinta, you're on the call, yes? Would you like to speak to this part?

SHINTA SATO: Hi. I have a little bit rewritten that part and added it to the next paragraph. What I was talking about was that adding that difference itself is a good thing among RSOs. But that is not measured in these metrics or any of this measurement way.

DUANE WESSELS: I see.

SHINTA SATO: I’ve added that.

DUANE WESSELS: Okay. I think I understand. You’re saying that we value the diversity and the different approaches but it’s not something that we’re going to measure with these metrics?

SHINTA SATO: Yes, true.

DUANE WESSELS: All right. Thank you. Any other comments?

PAUL HOFFMAN: I think that’s a great addition. I hadn’t thought of that but I could see that, without that, someone would look at some of what we were doing as a criticism of the diversity. I think that’s a very good addition.

DUANE WESSELS: Okay. Shinta, we can replace the above paragraph with the lower one, right?

SHINTA SATO: Yes.

DUANE WESSELS: Yeah, okay. Let me just make a note to do that. All right. I'm just noticing in the chat, Ken, that you said that you like the proposed text. But I don't remember what that was in reference to. Was that this section? That was in 4.8? Okay. All right. Thanks. I'll go back and look at that later. Any last comments about this section, 4.9, before we proceed? Okay. All right.

Now we're in the root server response latency section. We have this formula, here, which is our justification of how we came up with 250 milliseconds. It's based on the speed of light and whatnot. We have a comment from Kazunori, noting that in fiber, the speed of light is different, and that maybe instead of about 250 milliseconds the threshold for UDR response latency should be 400 milliseconds, which is closer to the metric in the gTLD guidebook, which is actually 500 milliseconds.

One thing I would note is that, in the case of the gTLD guidebook metrics, they're basing that on a 95th percentile, whereas what we're talking about is a median or a 50th percentile. I think that's one important thing to keep in mind.

If the group would like to increase the threshold, I'm okay with that. I do think at some point, though, we do have what Mr. Vixie, in the previous meeting, called a "marketing problem." I think that the higher that we make these thresholds, the more likely that people who are going to look at this are going to have questions about the thresholds. But if that's the consensus, I'm okay with it. Paul?

PAUL HOFFMAN: I don’t see a reason for updating this to deal with the slower speed of light in fiber because we have more than one vantage point. If we had a system with only one vantage point, we would need to worry about that. Since we have already said that the vantage points will be spaced around the world geographically … And even if we go to a topology methodology later, they still will be geographically around the world. And the fact that we have 20? It is literally impossible for a vantage point to need to go around the world to get to any particular instance. I think that the current wording and mechanism is probably sufficient, here.

RAY BELLIS: Sorry, Paul. I think I have to disagree with that. I think that is the argument isn’t based on the speed of light but on the speed of transit then we need to change the text.

DUANE WESSELS: Can you be more specific, Ray, about how you would change it?

RAY BELLIS: Ultimately, the speed of light to go around the planet twice is not, therefore, the criteria on which we’re making the measurements, is it?

DUANE WESSELS: Right.

PAUL HOFFMAN: Okay. I will grudgingly agree with Ray. I have no better way of coming up with a number. If we do speed of transit, we're going to be pulling more numbers out of our hat. I'm not sure how to do it. But I agree with you that what we are saying now, therefore, is not accurate.

DUANE WESSELS: All right. Ace, you have your hand up?

[ACE:] Yeah. I really don't think we should do the speed of light thing. I think that it will make us sound very confused or clueless to people who do transit stuff. It's kind of like the threshold is based upon the weight of an apple. It's unrelated to what network conditions actually are. For example, in many places in the world, 700 milliseconds is not an unheard of amount of time to be able to ping place on the far-side, or even relatively close, geographically. Tying it to some metric or some sort of system which isn't really related to what people see, I think, is nonsensical.

DUANE WESSELS: I guess one of the questions I have about this is that I feel like a lot of the Work Party participants expect there to be justifications for how we come up with thresholds.

RAY BELLIS: The figure sounds right. It's just that the rationale seems slightly bogus in retrospect.

DUANE WESSELS: Right. Would the Work Party be comfortable with just stating a number without a complicated justification or a formula-based justification? Is that something that we're willing to do? Paul, I know your hand was up before I asked the question so go ahead.

PAUL HOFFMAN: I was going to answer the question. I am willing to have a number. We don't even have to say we chose it arbitrarily. I think having a number like that without justification is acceptable. And if an individual asks us or an organization asks us, "Where did you get the number?" we can give a longer explanation of, "We pulled it out of the air because we tried doing other things and they did not lead to satisfactory results."

DUANE WESSELS: Does anyone else have an opinion on whether or not we need a justification? Assuming we don't need a justification, then what are people's thoughts on what this threshold should be? At our RSSAC meeting a few weeks ago, we settled on 250 but now we have a proposal for 400.

RAY BELLIS: To me, 400 sounds way too generous. I think I'm absolutely fine with 250. As Ken just said in comment, that's 50%, our median. I'm looking at my own stats at the moment and my F-root's current median latency, at least as far as the RIPE Atlas is concerned, is 12.9 milliseconds. 250 is so far beyond what we're looking at. Obviously, we do have individual nodes, the Atlas probes, that are far, far slower than that but they are far into the curve.

DUANE WESSELS: Ray, you're setting me up for what I wanted to talk about, as well, which was that I have been working on some analysis of the RIPE Atlas data. I think everyone sees the graph that I've put up, here. This is a little bit complicated so let me try to explain it. The RIPE Atlas system, there, is about 10,000 probes. I did a bunch of simulations showing that, if you chose a subset of the probes at random and then calculated the latency that you would get from that subset, what would that look like?

In this graph, the X-axis represents the number of probes that you would choose at random and the Y-axis is the latency. It's a little bit hard to see because there's a lot of alliance, here. But if you can pick out a single color, you can see that there's an upper curve and a lower curve. As you increase the number of probes in the simulation, they converge to a single point. But what this shows is that these points, here, are where you would choose 20 RIPE Atlas probes at random.

In the worst case, if you chose the 20 farthest away RIPE Atlas probes, the median latency is where this point is. And so, in all cases, we're below 250. In some cases, we're not really that far below 250. But in all cases, if you chose the worst RIPE Atlas probes, you're still below 250 as a median. Going with what Paul was saying, what we're recommending for the vantage points for this would definitely be located at places that are better connected than the worst RIPE Atlas probes, I would have to think.

[RAY BELLIS:] Well, we require them to be.

DUANE WESSELS: Yes. They're required to be at data centers with good connectivity, not at people's houses and in faraway places, right? Does this data make sense to everyone? I want to make sure that this is clear. Paul, is your hand up or is that from before? No. Shinta, go ahead.

SHINTA SATO: I still think that 250 is a little bit too short. I'd like to see 400 or something. This is the minimum requirement. It doesn't need to say, "This is the current state," but what we find acceptable is the value we are going to save, or the threshold. Operationally, I do often see the more than 250-millisecond response in our main servers. If the vantage point is set a little bit far from the name servers, this does happen. I'm very afraid of that.

DUANE WESSELS: Okay. Thanks. Ace?

[ACE:] Thank you. I would note that Shinta is most likely coming from Japan when doing the look-ups. I think that us claiming 250 milliseconds shows a very Westernized view of things. If you try and do these same look-ups from places like Africa, you get very different results. Latency is, from Africa, often 400 milliseconds, 450 milliseconds, 800 milliseconds. From lots of the Caribbean, bits of South America, Papua New Guinea. 250 milliseconds is the sort of latency people just don't see. I would think that we should be listening closely to what Shinta says and also taking additional measurements. The RIPE Atlas stuff is great but it does have an incredibly high European and North American bias.

DUANE WESSELS: Sure. But what I'm getting at here is that I think that this graph doesn't necessarily eliminate the bias. What I'm saying is that you can take the ten farthest away probes from a root and you still get a median latency that's well below the threshold, whatever those farthest-away probes are, whether they're in Europe, Africa, or whatever. Anyway. I'll stop belaboring that point.

Well, my feeling is that, since this is a change from what we talked about in the meeting with RSSAC, we can't really decide this right now. We'll have to take this to the list or take it back to the in-room discussions and leave this "undecided" for now, I guess.

All right. Go back to this one. Here, we're in the root server correctness section. Let's see, who was this comment from? Oh, this is from Paul. Paul, I forget where this text came from that you are highlighting, here. This is not something you originally wrote, I guess, right?

PAUL HOFFMAN: Correct. These were things that were added by Daniel.

DUANE WESSELS: Okay. The discussion here is about the purpose of this metric and whether or not it's designed to detect random modifications or whether all RRsets have equal importance in the detection of correctness failures. Unfortunately, Daniel’s not here, I think, still, to talk about this. Has anyone else had a chance to look at this and would like to make a comment about the highlighted section, here? All right. I'm not hearing anyone.

A little bit farther down in this same section, there’s text that talks about the two types of queries that are sent in these measurements, the expected positive and expected negative. In the expected positive case, right now, it just says, “This set of queries is taken from all authoritative RRsets from the root zone.”

Kazunori is suggesting that we should be more explicit, here, and describe all of the types of queries that this could be. We’ve had a discussion here on the side of the document. Personally, I think that, as written, it’s okay. It’s a little bit future-proof. For example, if different types of data are added to the root zone in the future, then we wouldn’t necessarily need to update this document. However, it is technically possible to list all of the types of authoritative data in the zone right now, if we wanted to. As Paul notes, it would be quite long. There are, I think, five, six, or seven different types but we could list them if we wanted to. Paul, did I miss anything in your concerns about this?

PAUL HOFFMAN: No, that’s correct. I thought listing by exclusion made it more readable. I think Kazunori’s request is that listing exhaustively might make it easier to understand.

DUANE WESSELS: Kazunori, did you want to say anything about this? Could you say why you felt it was important to list all of them specifically?

KAZUNORI FUJIWARA: Hello. I prefer listing all possible names and types. [inaudible]

DUANE WESSELS: I didn't quite get the last part. I didn't hear why it was important to you.

KAZUNORI FUJIWARA: I think [I tried to] [inaudible]. Without these delegations [on the case, the TLD list is nothing listable] [inaudible].

DUANE WESSELS: All right. I think we need to take this conversation, probably, to the e-mail, or we could have more comments in the document to make sure that I fully understand your concerns, Kazunori. For now, I think we'll leave this as unresolved and move on, if that's okay.

All right. So, still in the correctness section, we're getting into the part where there are descriptions of how you verify all of the different types of responses. I guess, to be fair, Paul, this list is probably not too dissimilar from what Kazunori is suggesting, which is that you list all of these different types or combinations of queries and names together.

PAUL HOFFMAN: Actually, Duane, that’s not true because we have combined all of the negative answers into one section.

DUANE WESSELS: Okay. But he was talking about the positive queries, right? Anyway. I think, Paul, that you and I talked about changing this a little bit so that it just references DNSSEC validation, which is similar to what Kazunori’s comment, here, is about, referring to RFC4035?

PAUL HOFFMAN: Yes. Actually, it’s interesting that Ray disagreed because the text that you’re seeing here, now, actually deals with Kazunori’s. That is, I made a bunch of changes and they were accepted. Ray, do take a look at the current text and see whether it deals with the use of DNSSEC in a way that you agree with.

RAY BELLIS: Yes. It actually does, more or less, except for the comments I just added onto the page below, where it looks like you’ve removed explicit mention of the RRSIG records. I think that, given what we’re describing in the packet contents, that the RRSIG should be mentioned.

DUANE WESSELS: Why do you feel like they need to be mentioned if we’re saying it’s validated with DNSSEC, since that’s the only way to validate, with DNSSEC? I ask this specifically because this is where Kazunori got caught up, earlier. I was trying to very carefully list all of the places we need

RRSIGs. I think Kazunori was correct in saying that this is making it more complicated than just saying "validate with DNSSEC," however you're going to do it.

RAY BELLIS: To be fair, I did feel that the old wording was a little excessive because you had a separate bullet point saying, "There will be this record," and then a separate bullet point saying, "There's the RRSIG for that record." I would be quite happy if it just said, "Assigned SOA record," for example. I'd like to just highlight that other people had seen that. Yeah. I would be happy if it said, for the last section, "Assigned SOA record for the root."

DUANE WESSELS: Okay. I'm sure we could agree, but saying that it contains a signed SOA record doesn't say that we're validating the signature.

RAY BELLIS: No. I did say in my comment that there should be something at the top that basically says that all signatures are validated. "All signatures must be validated."

DUANE WESSELS: I can absolutely see that is to get rid of the "and that [RRSIG] was validated with DNSSEC," get rid of all of those, put something at the top, and just say which things are signed.

RAY BELLIS: Yes.

DUANE WESSELS: I can see that. Kazunori, does that also still satisfy your desire for making this simpler?

KAZUNORI FUJIWARA: In my opinion, it is the correct specification we've had, I think. That's all.

DUANE WESSELS: Okay. Let me make a run at this and I can do that later today. Ray, you can see whether I hit that more correctly, and then we can maybe ask on the list again, if the list is then still correct.

RAY BELLIS: Yes, please. Just hit me up when you've done that work, at some point, and I'll look over it. Having been through the process myself, writing that server, is code you've referred to in your own comments. It's actually quite useful to know, when you're looking at answers, to say which RRSIGs are really supposed to be there or not. I think it's simpler to just say "it's a signed record" than have somewhere else say "signed records must be validated."

DUANE WESSELS: All right. Great. Thanks, guys. We're now in the root server publication latency section. There's a comment from Daniel talking about a couple of things. Currently, the text here says, "This processing system needs to know, approximately, when new zones are published. This could be done by looking at notify messages or it could be done this other way, by examining a collection of SOA serial responses from all of the RSOs."

A conversation that I had with Daniel about this talked about the notify part a lot, I guess. I put this in here. I put the notify part in there. I forget where it came from, originally. That's one way to do it. It doesn't have to be done that way. I get the sense that Daniel would prefer to not mention "notify" at all, which is fine with me. It doesn't have to be that way. We could specify just to look at the SOA responses collectively.

Daniel also has some comment about looking at … I'm not sure if it's a different way of determining when the zone is published or not. It's a little bit hard for me to follow his comment. But I think that the way that the document currently described getting the SOA serial numbers and detecting the new zones is sufficient for someone implementing these to go and do that. I don't think we need to be even more specific than that. Are there comments from anyone else about this issue?

PAUL HOFFMAN: I like getting rid of anything about "notify," since that's not something that exists currently. It is something to be added but that capability does not exist currently. I think it might be confusing to folks.

DUANE WESSELS: Okay. I get the sense that Daniel would agree with that, so we can make that change. All right. That takes us to the end of the RSO metrics. Within the RSS metrics, Daniel made a comment a couple of times that there's this inconsistency in that sometimes we use the best K subset of data to calculate a metric and sometimes we don’t. He thinks that it should be consistent, except in the case of correctness.

My opinion is that that doesn't always make sense, to do it that way. But I kind of see where he’s coming from, that people could be confused. This is one place where he mentioned this. I think it’s mentioned later, as well. I'm just going to save that discussion for later.

Ray, you have this comment, which has been here for a couple of weeks, about how, in this particular example, it would indicate a problem with a vantage point. We still have it as a to-do item, adding text that describes the need to detect when vantage points are having problems and eliminate them as false positives, as you say. That’s still work to be done.

RAY BELLIS: Yeah. I don’t have anything further to add to that, at the moment.

DUANE WESSELS: All right. In the RSS response latency metric, again, we have a discussion here that focuses on the K-on-N issue. This is from Shinta. He notes that there’s no reason to expect that the distribution of latencies would have a linear relationship. And so, it doesn't make sense to use this formula, here.

Getting back to our previous discussion about the need for justifying our thresholds, maybe, here, we just eliminate this and come up with a number again. What we came up with from the meeting with RSSAC was 150 milliseconds UDP latency. Shinta, are you still okay with this? Since you were proposing increasing the RSO threshold, would you propose increasing this one as well, or would you keep it at 150? Shinta, I saw your name pop to the top of the participant’s list but I didn’t hear anything from you. I'm not sure if the audio actually worked.

SHINTA SATO: Hello?

DUANE WESSELS: Yeah, I can hear you now.

SHINTA SATO: Okay. Sorry. I was muted by my microphone. Okay. If the RSO metrics threshold has changed, I think this should be changed as well in the same factor, or something like.

DUANE WESSELS: Okay. Thank you. Ace?

[ACE:] Unsurprisingly, I agree with Shinta.

DUANE WESSELS: All right. Thank you. That is future work, then, to settle on the metrics. I assume that there were no objections to removing a formula-based justification for this threshold? Okay.

In the RSS correctness section, there’s a comment from Daniel, again about how, in this case, we would not apply the K-out-of-N model, which I think everyone is in agreement with. I'm not sure that the added text is necessary, though. We’ll take that discussion up with Daniel, probably offline.

In the publication latency metric, Daniel is proposing to use the K-lowest latencies. So, similar to what was done for response latency, he would say to do that for publication latency. Personally, I don't think that makes sense but I would like to hear what other people think.

RAY BELLIS: There’s maybe some merit in that. In a hypothetical situation where a root operator might miss a serial update for whatever reason, if it’s just the one, strictly speaking, there’s no harm done. [The root system won’t affect it], effectively.

DUANE WESSELS: Right. The threshold is based on the median of all of them so, certainly, one RSO being very late wouldn’t affect the median, I wouldn’t think. You would have to have a lot of root servers do poorly to affect the median.

RAY BELLIS: Yes. If this is for the system as a whole and it’s just a median then yes, that would be fine.

DUANE WESSELS: All right.

PAUL HOFFMAN: I fully agree. I can't imagine a situation where even half of the root server operators will have missed a 30-minute window. And if that's the case, I don't think we should raise the threshold. Remember, people will be looking at this data not just to look at thresholds. They'll be analyzing it. If we start seeing months where that happens, that indicates a problem with the root zone maintainer. Apologies to Duane. And therefore, something needs to be done. Remember, the thresholds are not the only thing that will be looking at these measurements.

DUANE WESSELS: Yeah. And then, lastly, there's a suggestion from Shinta to make the threshold, here, the same as the threshold for an individual RSO, which is one hour. Previously, we had agreed that, to make them different, this one was a little bit more aggressive, at 30 minutes. Ace, your hand is up. Go ahead.

[ACE:] Sorry. A quick question. Can somebody remind me what the polling interval is going to be?

DUANE WESSELS: Five minutes.

[ACE:] Okay. Never mind, then. I'm happy.

DUANE WESSELS: Okay. Cool. I have some sympathy for Shinta's suggestion because, really, this is just a straight aggregation of all of the individual RSOs' data. Assuming we reject Daniel's idea to use the lowest ones, then it would kind of make sense to have the thresholds be the same. Any comments about this before we move on? All right.

There was one meta-question that I wanted to talk about. Unfortunately, we only have about five minutes left on the call. At the top of the document, there are to-do list items. One thing that we've talked about a number of times, which still is unsettled, is this number four. We have a terminology problem.

A lot of times, we talk about "RSO" when we really mean "root server identity." We talk about "RSO latency," which really should be "root server identity latency." We kind of need to agree how this is going to be represented in the document. One of the reasons is because I think we don't want to confuse our readers on what we mean when we say "RSO." We want to be very clear. And to be honest, our current usage of "RSO" in this document is probably not as clear as it should be.

It's nice to be able to use acronyms, to be able to say "RSO." We could say "RSI," which sounds a little bit strange. We could say just "RS," for "root server," which also sounds strange. But I would definitely like to have people's input on this question. Paul?

PAUL HOFFMAN: I agree that we need to differentiate. I have a problem with “root server identity” because root servers have two identities – an IPv4 address and IPv6 address – which we can collapse in general talk but we can’t collapse here, in the metrics discussion, because we’re measuring them differently.

DUANE WESSELS: Yes, that is true. Really, this is awkward because what I think we want to say here is a letter. But for various reasons, RSSAC is trying to not talk about letters. But that really is what we want to capture, here.

PAUL HOFFMAN: Exactly.

DUANE WESSELS: Andrew, you’re on the call. You haven't started on the terminology updates yet, is that correct?

ANDREW MCCONACHIE: I have started updating RSSAC026 in the … The old word, or the old term, was “root server identity.” Based on what I’ve seen in this document, I updated RSSAC026 to just use “root server.” I can change RSSAC026 again. This is RSSAC2062. It’s just going to take whatever comes out of this document. I had no strong opinions on anything, as long as …

DUANE WESSELS: I don't think you should go and change things in RSSAC 026 just because of this, yet. That would be premature, I think. But we do want all of these things to agree with each other, right? So we need to coordinate that. Paul, is your hand up again?

PAUL HOFFMAN: Yes. With respect to RSSAC026, I think we really do need to get these two in line because I don't think that the metrics document is going to be the last document that has this issue.

DUANE WESSELS: Andrew, is there a definition of "root server identity" in RSSAC026, separate from just "root server"? There's not, right?

ANDREW MCCONACHIE: I think we changed the term. I think in RSSAC026, "root server identity" is what's being used. And then, in RSSAC026 v.2, we changed it to just "root server," if I remember correctly. But I'd have to go look.

DUANE WESSELS: Okay. All right. Thanks. Russ, your hand is up?

RUSS MUNDY: Yes. [inaudible] is identify a specific lump of time to use in a caucus meeting at ICANN66 to talk about this terminology challenge in combination with not just this document but with 026. And like Paul says, others will come along later. Perhaps we should try to, maybe, tackle that issue before we actually get into the next review of our metrics document.

DUANE WESSELS: Okay. That's a good suggestion. Thanks. Ray?

RAY BELLIS: I was just wondering whether we could consider "root server address," as opposed to "identity"? For example, in [F area], 19255241 is its identity for IPv4, with a similar address for IPv6.

DUANE WESSELS: Yeah. At this point, all ideas are still on the table. I think it's a matter of finding a term that resonates with us, that we like. You would have to imagine that if we said "root server address" then, do you want to talk about "root server address latency"? That's the question for me. Ace?

[ACE:] Thank you. I know that RSSAC is trying to get away from the term "letter," but I think that we're doing all sorts of weird machinations and jumping through hoops to avoid it. "Letter" is what is generally used out in the world. It's what people refer to them as. You know, "The root server letters, blah, blah, blah." I think that the very fact that we're going through these weird addresses, identities, or other things, is kind of evidence that at least in some sets of discussions "letter" probably is the right term because it's the understood and known term.

DUANE WESSELS: Yeah. I think that's right. But one of the reasons that we got to this point was because the use of "letters" was almost a little bit too prevalent. People were saying things like, "K-root" when they really meant "RIPE." We've kind of swung back to the other extreme, at this point, I think.

[ACE:] That is true. We're using that. But here, we're actually referring, I think, to the root servers themselves, not the organizations. I think that "letter" is appropriate. I'm sorry for jumping the queue.

DUANE WESSELS: No, you're right. We'll make sure that we can talk about this at the upcoming caucus meetings, as Russ suggested. I'm sorry we went over time a little bit, everyone. Any last things before we wrap it up?

There are upcoming caucus meetings. There's one at the ICANN meeting in Montréal. Everyone, be sure to look for remote participation information on those so that you can stay abreast of what we talk about. And then, we have another one a few weeks later at the IETF meeting in Singapore, where we'll probably cover a lot of the same material. All right. Thank you, everyone, for participating. I appreciate all of the feedback. We'll talk to you soon. Bye.

[END OF TRANSCRIPTION]