

---

MONTREAL – RSSAC Work Session 3  
Sunday, November 3, 2019 – 13:30 to 15:00 EDT  
ICANN66 | Montréal, Canada

STEVE SHENG: Duane, this is Steve. In the caucus meeting, we also have an agenda topic on the Metrics, but that meeting has other items, so...

FRED BAKER: Okay, so the topic for the next 90 minutes is Metrics, and Duane and Russ, let me turn it over to you.

DUANE WESSELS: Alright, thanks everyone. So this is Duane Wessels, co-chair of the Metrics Work Party and Russ is sitting next to me, I'm sure he'll introduce himself later. I'll give just a brief update I guess since the last time a lot of us met to talk about this was in the Reston Workshop a month or so ago, and there's been a couple of Work Party meetings since then, where we've talked through a lot of this.

Quite a few things are changing and evolving. Our goal was to be at a point today where we had a final draft but we're not quite there yet, we still have five or six outstanding things that we need go through, that's what we'll do today and try to get through all of them. Steve is running the document up in the Zoom Room or Ozan is. Ozan, if you can scroll down until you see some comments on the right side, that's where we'll go here.

---

**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.**

---

FRED BAKER: While he's scrolling, let me ask a logistical question. Let's assume we don't finish today and let's get through as much as we can of course, but then we do have calls in the future or are we meeting in Singapore? What's happening?

DUANE WESSELS: There is of course and RSSAC caucus meeting in Singapore, I don't believe that -- we did have regularly scheduled Work Party calls but I believe that there was end date which we're at now. I don't think people's calendars have more RSSAC Metrics meetings on them. We can address that I guess. Our goal was to move to the next phase for this document, which is to get caucus approval and then RSSAC approval. We're almost there.

FRED BAKER: Yeah, let's see how much we can get through today.

DUANE WESSELS: The first point in the document that we want to bring to everyone's attention is in Section 4.8, the title of this section is Unexpected Results and Paul Hoffman has graciously added some new text, describing the situations and reasons that someone who's operating these metrics may want to or may need to exclude vantage points from the set of metric because the vantage point is offline, disconnected or otherwise providing bad measurements.

---

This text looks good to me, I don't have any problems with it but I wanted to bring it to everyone's attention, to review it. This also addresses some comments that are a little bit later in the document, assuming we like what's here, then we can resolve those other later comments. Paul, is there anything you wanted to say about this text?

PAUL HOFFMAN:

Yeah, Duane, I wanted to say this to you last week, as I wrote this text, I discovered something that's unfortunately a vocabulary issue, another one, which is, we don't actually consistently talk -- we talk consistently about the vantage points, we don't consistently talk about the thing that isn't the vantage points that is doing something, that is turning measurements into metrics, which may or may not be the same entity, which is looking at the metrics and seeing whether any particular metric does not meet a threshold, which may or may not be the same entity which is telling the world, this threshold was not met, they're all met or not and I don't know if we want to say it's one entity?

If we want to have different names but the wording in here, it's not like I'm clear in this edition either. I just wanted to bring that up but we really don't have a consistent name for the collector.

DUANE WESSELS:

Maybe that deserves a little enter in the terminology section, to sort of get us past that. I guess it's always been my assumption that the party

---

doing the collecting is also the one producing reports, which essentially say whether the thresholds are met.

PAUL HOFFMAN:

I didn't want to do that because I didn't know if people agreed that that was true, that this document has just two parties, I'm sorry three parties, the RSO's, the vantage points and the dollar sign X. If people agree on that, then we can come up with whatever name for dollar sign X, but if people don't agree, for example if they want to have the reporter be a specific group that is not the one who is doing the mechanics of creating reports, then we need to come up with dollar X and dollar Y.

RUSS MUNDY:

I think you raise a good point, Paul. I'm wondering if we could change the term party to activity, so if in fact it was a different activity that happened to be doing the collecting then was doing the aggregating and reporting, it wouldn't particularly matter for our semantics in this document, I don't think. I haven't gone looking to see if it would fit everywhere but the general idea is that they could be one or it could be separate. We don't have to worry about it if we say activity.

PAUL HOFFMAN:

But then we would need say which activity are we talking about, are we talking about the aggregating activity or the reporting activity? You're using activity, you're kicking the can down the line because we

---

couldn't just say the activity, we would have to say which one. Anyways, Duane, I think that's an open issue, I don't think its necessarily contentious but it was something that I saw when I was writing these paragraphs.

DUANE WESSELS: Okay, thanks.

RUSS MUNDY: Is there any views that RSSAC or caucus folks want to comment on? Just one entity or we want to leave it open so there could be two separate entities? Open. Anybody else? Okay, we'll look at seeing what the right way to do this. A slight preference for leaving it open.

DUANE WESSELS: Let's scroll down farther. Here's a fun topic, in Section 5.2 we have the definition of root sever response latency and when we left our meeting a month ago at the Reston Workshop, we had agreement that the threshold here should be 250 milliseconds for UDP and 500 for TCP. Within the last couple of weeks there were some comments and we had a discussion that -- first of all, we had a justification in here that was based on speed of light around the earth twice or something like and there was a formula and it was pointed out that if you wanted to do this properly, you need to account for the speed of light in fiber not in open space and so that changes you calculated value from something like 250 to more like 400.

---

We talked about this on the call last week and came to this consensus that maybe we don't need the formula to justify our choice here but we still do have the open question of, what should this threshold be? We had agreement on 250 but now there are advocates for a higher number of 400 milliseconds. Also, part of that discussion was that 400 milliseconds gets us closer to the gTLD metrics thresholds, which are 500 milliseconds, however in the case of those gTLD thresholds, it's a 95<sup>th</sup> percentile whereas this is a 50<sup>th</sup> percentile. I didn't see whose hand went up first. Michael?

MICHAEL CASADEVALL: Couple points here. The first point I'm going to make is that the speed of light, when travelling through copper is two thirds C, so you have to account for that or other mediums.

DUANE WESSELS: We're not going to use that as a justification. We're just going to choose a number.

MICHAEL CASADEVALL: Right, fair enough. The other point is that, I'm concerned about us having a generic threshold here because of network conditions, unless we're monitoring one point with a known route to all the roots that stay static, this number could fluctuate way too much. I have concerns that 250 is too low, I'd be happier with 500. I worry unless we have a clear of how we're actually gathering this number consistently,

---

I don't know if it's going to be that valuable because of where you're measuring.

DUANE WESSELS: The way we're gathering this number is written in the document, we're spent a lot of time talking about that, I think that's sort of well described. The really question is the threshold. Fred, did you get a chance to say what you wanted to say?

FRED BAKER: Well, I was wondering if you had had a chance to look at theRIPEATLAS data and qualify the number there?

DUANE WESSELS: I actually briefly presented some of this at our Work Party call last week. One of the things that I did with the RIPE ATLAS data was some simulations, there is 10,000 ATLAS probes and I showed that if you took a root server operator and picked the farthest away probes or the 20 worst performing probes, on a typical day you're still under 250 milliseconds latency threshold.

To me, 250 is reasonable. I understand what people are saying, there are reasons why it might not always be 250 and maybe we should give ourselves a little more room but I also did emphasize that we're talking about a median value, so we're talking about only 50 percent of the responses have to be less than 250 milliseconds. 50 percent can

---

be 10 seconds, not really because that's above our timeout but they can be very high and you can still meet threshold.

I think the current data supports 250 reasonably well but I don't want to discount the input from the other people who feel differently.

UNKNOWN SPEAKER:

I think even looking at the RIPE ATLAS probes is not as realistic as test. I think looking at the anchors, which are more suitable or more equivalent to what the root server system clients are going to be, I think you'll find even better round trip times than basement probes. 250 seems more than high enough to me. It's a global system, it's any cast, they should be everywhere.

BRAD VERD:

I feel 250 is a generous number in my opinion. I'm fine leaving it as is. I guess I feel like I need to put a qualifier out there. I feel like I've said this numerous times but it doesn't seem to stick and maybe this is my perception, so please tell me if I'm wrong but I feel like these numbers that are being worked out here in this Metric Work Party, the biggest concern of these numbers, this is going to be some bible that doesn't get changed, doesn't get reevaluated for how you get rid of a root server operator and that's not the case here.

This is our stab, RSSAC kind of defining what metrics to what the root server system -- it's our first stab at creating thresholds on these metrics. I think what will happen is that and what has been stated is

---

that the output of this Work Party, would be handed over to the GWG as input into whatever they come up with to solve the PMMF and what not.

I imagine this will change when it goes through the community and all the other stuff, so this isn't going to be the triggering event, defined here. I think once you -- I think we need to put a line in the sand and I think 250 is a very good line in the sand. I think on all of these metrics, I think we need to come back after six, twelve months and reevaluate and I think we'll know more after we collect this data, after we view the data in real world life and see what works and what doesn't work and start to understand what the anomalies might be. That's just a disclaimer.

DUANE WESSELS:

Paul?

PAUL HOFFMAN:

Most of what Brad said but with one change, when he said we're going to have to look at it in six month; the GWG also has the option of saying, thank you RSSAC for that number, we're going to do an independent study or we're going to ask for research or whatever. Yes, we need to draw a line in the sand and by the way, I totally agree with Brad, 250 for a median seems very, very generous, given that there are going to be 20 collectors, sorry 20 vantage points, I need to get my own vocabulary right, 20 vantage points throughout the world and such like that. I

---

f half of those, even if you are a very, very localized root server, you're going to have a bunch of them in your neighborhood. If it turns out that this totally wrong, we can revise it later but I would not try to say, we're going to keep revising it up to be nice, I would say, let's pick a number that we believe is a reasonable value for what a root server operator should be doing in aggregate over a large collection.

UNKNOWN SPEAKER:

Just to second Brad, the thing is that if the number seems too high or too low, I think we have to consider it's a median and its vantage point is well connected. If RSO's are not afraid of this number, it's fine for a first shot.

WES HARDAKER:

It sounds to me like there is sort of consensus, A where everybody was saying is 250 is fine, so let's run with that. I guess to push back a little bit, not really push back but to restate it in a different way, we really have no idea how this document is going to be used. We are defining it today, we may end up measuring it today, whether something else just refers to it and says, that's the golden stamp of they're going to completely rewrite it we have no idea.

We don't know if it's going to be the GWG, we don't know if it's going to be some element of 37 in the future, we don't know. We are doing this as preliminary work to help feed that, to say we've already done a lot of thought, this is what we believe the system today, not five years from now, today should be able to do and to tolerate.

---

Now, what will happen immediately is if we do stand up these numbers and we start monitoring these numbers and if somebody doesn't meet these numbers, there will be communication within the ICANN community about what to do about it and there won't be a place conversation to be held until the rest of everything else is stood. I'm sure the Board and CTO and lots of people will get an earful about it, hopefully that won't happen and hopefully the numbers will work out, if everybody believes they can meet 250, then I think we're good. Nobody has said 250 is bad, let's go with that.

DUANE WESSELS:

One of my frustrations, we have non overlapping sets of people in our different meetings about this topic. Here everyone in the room can probably agree on 250 but then we have a call with different people, the conversation goes a different way and people have different opinions. It's a little bit hard to get consensus in these cases. I don't know how to solve that problem.

RUSS MUNDY:

A way we can do it, is to reach a conclusion here, part of the conclusion being that we will send something, a statement to this effect on this item to the caucus mailing list with a defined time period for anybody to raise any specific objections and state the reason why and if they don't, then it's solved and we declare because this has been one of the most contentious things we've had. We just declare it, whatever we agree too today, if we don't get any disagreement by the

---

caucus meeting in Singapore. It's a path forward, it draws a line in the sand.

DUANE WESSELS: Right, but still someone can -- I feel like that's already happened. There are comments on the document and they've stated their reasons why it should be 400. We have strongly differing opinions here.

RUSS MUNDY: Have we asked the Work Party as a whole, that -- I'm trying to remember the last time we specifically asked the question. I'll have to go back and look and see if have asked it specifically. Previously we've kind of said, we're moving forward, put your comments in the document without really flagging explicit issues. This is one really explicit issue that we have. That's my suggestion is we just draw a line in the sand, for whatever we decide here and then if we don't have anything sound by the Singapore Caucus meeting, if we do have input by then, we discuss and make the final decision at the Singapore caucus meeting.

DUANE WESSELS: Let me ask a hypothetical, would we be okay as a Work Party, as RSSAC, if there was a dissention to this, somebody use the descent option of this document to say that they didn't agree with these thresholds? Is that a reasonable thing to do in this case?

---

RUSS MUNDY: I would prefer not to have descents but it's part of the process and its part of the system that we have in overall procedures so, then from my personal opinion, it would be okay if someone felt [inaudible] about it.

WES HARDAKER: I was but I think I'd just derail further so I'll pass.

PAUL HOFFMAN: I am hesitant because, Russ, you said actually two different things in two different joining sentences about a minute ago. One was, did we ask for something specific on this one and you also said, maybe we just draw a line in the sand on everything. I believe a saner proposition would be the later, that is we go through, we take whatever the current consensus is, we have a clear document and then we say, are there any issues with anything in specific or do people want to say that they like the whole document?

I think if we keep doing this piecemeal, it allows people to open up piecemeal things. I believe if we have a whole document, because some of these parts are interrelated, this part is in fact interrelated to the RSS thresholds and such. I think we're better off instead of saying, this one was contentious, let's talk about more is to forge forwards and say, where are the objections now that someone can see the entire document. Thank you.

---

**DUANE WESSELS:** I'm going to move on in the interest of time. Just a little farther down, in Section 5.3 there is a new paragraph that talks about correctness and reasons why responses may be correct. For example, because it was incorrectly transmitted a root server operator or due to some kind of attack in the network on the path. Then it references the earlier advice about spoofing protections and dealing with unexpected results. I think this is a good addition. I was waiting for Kazunori Fujiwara to give his thumbs up since this was something he wanted to see in the document. Any comments about that? Okay, good.

Going down a little bit farther, still in the correctness section, this has been a little bit fluid here. Ozan, can you go back up to the table, there is a table in contention here. We've sort of agreed that we need to be specific in listing the names and types of queries that will be made for correctness checks, I put this table in here, Paul is advocating for not a table, for just putting in line and in text. If anyone has an opinion one way or the other, now would be a good time to say, otherwise I would probably defer to Paul's preference.

**PAUL HOFFMAN:** Not speaking about my visual preferences, I probably should have done a separate note. You left out .NS and I just want to make sure that you actually feel like you left it out by accident.

**DUANE WESSELS:** I will agree to that omission. Well make that change a little bit later. Let's scroll down some more. The correctness check is quite long,

---

we're still in this section. Paul has highlighted some bullet points about how you do validation and checking. I think, Paul, you're just looking for another set of eyes on here, to make sure that you got the wording right, yeah?

PAUL HOFFMAN:

Yeah, I'm pretty sure I got it wrong. For anybody who's had to describe negative answers, DNSSEC answers, please take a look at this because the question of NSAC covering and such like that, Matt, I'm sort of looking at you since you're one of the authors of 4035, but I just had a really hard time saying that in the response, you see blah, where blah means a verifiable negative. I put some words in there but they just feel wrong. But I think everyone knows what I'm trying to say but that doesn't mean that the next people will understand.

DUANE WESSELS:

Thank you, Paul. Ozan, let's go all the way down to Section 6, the very start of Section 6. As a little bit of reminder, previously we came up with this formula and the description around using K of N as the basis for some of the metrics. For example, in determining how many root servers needed to be operational and so on. Daniel has a comment here that he's sort of suggesting that the idea of relying on K of N, maybe we should use that more broadly in all of the RSS related metrics, not only the availability metric.

For example, in RSS availability we definitely use that K of N concept and we also use it in the response latency. The idea was to take the

---

latencies from the K best servers in each interval but we do not use K of N for correctness and Daniel is not suggesting that we do. We also do not use it for the publication latency metric, which I think is where you are suggesting that we do.

DANIEL MIGAULT: For correctness, we do it but with K equal N.

DUANE WESSELS: That's too confusing. We can't do that. I guess the question to the group is, do we want to consider this K of N is a general property of all the RSS metrics or only apply it to specific metrics where it makes sense? Does that capture proposal?

DANIEL MIGAULT: Yeah.

DUANE WESSELS: My personal recommendation is I think it's sensible as is. It feels a little bit strange to me to apply it to the publication latency, especially given that really the only reason we have publication latency in the RSS metrics is because its omission was noted before. It's not a super critical metric, I guess. I think I saw a hand go up over here.

---

BRAD VERD: I don't see a need to use it everywhere unless there is a problem everywhere. This started out as an availability equation that -- it was design equation that was used and the example was given for aircraft type of thing and then we've kind of stretched it a little bit to try to fit it in to our scenario. I don't think just because you use it in one place, you got to use it everywhere unless we're trying to solve something. Is there something that's not working with what we have for the ones that aren't using it?

DUANE WESSELS: No, I don't think it's like that, I don't think there's something missing or something that's not working. I think the question a naïve reader might have is, why did you use K of N in one place and not use it in all the places. Really, we're using it in two places. Why did you use in two places but not in the other places?

BRAD VERD: It would be one thing if I used it in one part of availability and not the other part of availability but we're not doing that here. I don't see where a reader would get confused, I don't, but that's myself.

PAUL HOFFMAN: This is really a question for Daniel. The K of N is actually not introduced here in Section 6, it's introduced in RSO availability. To me, either we do it in all of the RSO metrics, which I think really makes no sense or not. I don't see why it is brought up here. Why do you

---

think that this was part of the RSS, when in fact K of N was specifically done for the RSO metric?

DANIEL MIGAULT:

Don't we do that in the RSS? I think we do that.

DUANE WESSELS:

The first that this idea of K of N is introduced is actually in Section 4.9, Section 4 is general things about the metrics, 4.9 we have the formula and a little graph that shows how K changes when N changes and things like that. It does appear very early in the document. We use it in Section 5.1, which is RSS Availability, as way to get to the threshold value but not as a way to perform the measurements and metrics, it's only used to get to the threshold. Then in the RSS sections we use it in RSS availability and in RSS latency.

LARS-JOHAN LIMAN:

I have an idea for a divider here between using it and not using it, I don't know if it applies because I don't have all the metrics at the top of my head but could it be applied for things that are quantitative but not for things that are qualitative? Saying the quality of the root zone, meaning that we all have to provide, the t updates time for root zone is a matter of quality, whereas the availability of servers is a matter quantity. Does that work as a divider for people? My hope is that, that's already the case, so that we don't have to change anything.

---

DUANE WESSELS: I think you could argue that is already the case and I kind of like that but I'm also at the same time, a little bit afraid of having to write text to describe what we mean by qualitatively, I think that's a time suck, I don't know.

UNKNOWN SPEAKER: My position is that I'm fine with the consensus, I think it's better that we have one way to express the RSS metrics but I'm fine we're using different definitions.

DUANE WESSELS: That's good, we can take that as resolved. Ozan, if you scroll down to Section 6.1, there's a comment from Ray, which I think this is now addressed due to the recent additions from Paul about vantage point availability, connectivity and things like that. I think that's okay. Under RSS Response Latency, Section 6.2, there was some notes here that if we were going to change the thresholds for RSO Latency, then they should probably be changed here as well. I guess we'll keep that here for now and resolve them both at the same time. In RSS Correctness, Daniel, this is another comment from you. Is this along the same lines as before about using K of N, I think? Okay. So, I think that --

---

DANIEL MIGAULT: From what I recall, that was the only metric where we were not using specifically K of N -- the way to describe that metric would have been to say, we raised the number of K from eight to N.

DUANE WESSELS: Then in the Publication Latency Section, again there's a comment from you, Daniel, about how we do not use the lowest K of N, which I think that is the right thing to do but I'm willing to be convinced otherwise. Then additionally in this section, there's a note about the threshold.

At the end of the RSSAC Workshop in Reston, we had agreed that for the RSO metrics, Publication Latency, the threshold would be one hour, however for the RSS Publication Latency, it should be lower, we should ascribe to a higher standard and it should be 30 minutes but there is an argument that, why are they different when they're really measuring the same thing. It's just a bigger aggregation of the exact same data set from all the other operators, so why should it be any different? I'm a little bit sympathetic to that point of view as well.

PAUL HOFFMAN: For exactly the reason you just said, it's a bigger aggregation. That is, we do not expect someone to be meeting a minimum threshold for an entire month.

DUANE WESSELS: Or you do not expect everyone to be meeting the minimum threshold?

---

PAUL HOFFMAN: I'm sorry, you don't expect everyone -- let me be a positive. You expect some root server operators to act better than others. You expect some root server operators to work as hard as they can to act in the best interest of the root server system, therefore if we came up with a threshold that even the very best root server operator -- let's say half of the best root server operators, the ones who are doing the best, could only barely meet the threshold, then we made the threshold wrong and in this case, we didn't.

DUANE WESSELS: Any other comments on that? Shinta, did you want to say something about this? Are you fine with 30 minutes?

SHINTA SATO: I just wanted to mention while it is up, I couldn't find the reason.

DUANE WESSELS: I see what you were saying. There was no rationale for why it's exactly half of the other one, like how did we come up with 30 minutes?

PAUL HOFFMAN: Duane, maybe add a sentence of rationale?

---

DUANE WESSELS: Yeah, but I feel like we're being inconsistent because we deleted the rationale for the Response Latency Cases, that's a little bit frustrating. Either we have rationales for everything or we don't need to have rationales for number that we chose.

WES HARDAKER: I don't think you have to have a rationale for everything. When you can come up with one, great but sometimes you're going to end up a value that there was a lot of conversation around and we okayed it as a value, even though nobody came up with a perfect logical reason behind, it just maybe how the current system works. In this case, you actually have a real reason why you're picking the value you do, but as long as I have the mic, it should probably be one hour plus five minutes because you really want to two plus, if you're really trying to wait to let them have two chances to pull the zone, it's not instantiations, so you might say two plus a few minutes.

DUANE WESSELS: Right now, we're in the RSS section and I think the consensus was to put this back to 30 minutes, so I guess you're saying 35 minutes and then the other one an hour and five minutes. Okay, that's fine.

That's probably most of the bulk of the things we still need to do. Ozan, if you scroll down to -- go back up a little bit. There's a list of recommendations here. This has not been changing recently, at least not that I can remember, so we're pretty good. However, I do want to call your attention to Recommendation 2, which is just down a little

---

bit lower. Recommendation 2 is the one that says, that the RSSAC should do a future investigation into a better way to distribute vantage points by network topology instead of geographic location. In mind that's a future work item.

Then if you scroll down to Section 9, we have a new section called Possible Future Work, which has some other things which are kind of along the same lines. Currently they're not written as recommendations. The Future Work Items are, to consider -- and these are things that we discussed at the last workshop, to consider self-reporting for more coverage and eliminate uncertainty of components not under and operators' control.

Number 2, to create a reference dataset. Number 3, to explore financial aspects of increased accountability and how it might relate to these metrics. Number 4, to create a document that advises yet to be created functions on how they should interpret and act on data from this measurement system. My question I guess to the group is, do we want to make these stronger recommendations? Are we happy with the current recommendation to as a very specific recommendation but leaving these are more vague or do we want to make them all the same or leave them as is?

Steve and I had a discussion about this right before hand and, Steve, I think you said that if it's written as recommendation then there's a pretty strong expectation that the RSSAC would actually follow up and do that work, it would be tracked in some way, right?

STEVE SHENG:

Yes. My sense is, whatever the work party decide that the RSSAC definitely going to do follow up work, we put in the recommendation and if there's no agreement that is happening sometime very soon, we just put it in the possible of future work section. That means give the RSSAC more leeway in deciding the work, whether to undertake that and the priority of that. That's my sense. With that particular, I would think recommendation 2 should be removed and that bullet put into possible future work section, that will be my thought on that. Thanks.

RUSS MUNDY:

I think had enough discussions in the work party about a variety of different things that are really kind of collected in Section 9 and we had not gone through the moderately rigorous effort of saying, are all of these things, things we want to as RSSAC explicitly work on in the future? Of course, now, the RSSAC things are things RSSAC caucus will actually be doing and in terms of those folks that were pushing quite hard for having a future approach for vantage point, consider topology in addition or instead of geographic location, it would in some ways be lowering the emphasis on that particular point.

I wanted to at least note that and I don't even remember who it was that was pushing vigorously for that. If we just take that, drop recommendation 2, make 3 number 2 and include in Section 9, does anyone who was pushing hard for the topology-based location object to that? If we don't hear an objection, we'll move it.

DANIEL MIGAULT:

I don't think it's an objection but when I'm writing [inaudible] works as a recommendation that we should revise things, yeah, I think the document should be clear that it's a starting point. Something I'd like to see maybe in the future work or somewhere is that it's a non [inaudible] approach and we don't necessarily have a clear understanding of the metrics behind.

I think that should be also stated that there is maybe some future work. If the topology is moving, what is the impact and how can we check that we are actually measuring what we expect to? This kind of things should be -- we need to cautious around those in this document and I think there is a place for future work in this direction. I don't think it's something that should block the document from moving forward.

RUSS MUNDY:

So, we'll go ahead and move it into that section. I think one of the points that you raised is, we need to be sure we have words in Section 9 that address the fact that the current -- although there's words in Section 2 and Section 4, generally to the affect that this is a first effort, a version 1, it doesn't hurt to make sure we have something at that nature in this Section 9 too.

---

BRAD VERD: Just for clarification, you keep saying recommendation 2 and I keep going back to 7.2 in the recommendations, and I'm just trying to make sure that's what you're referring to, is that correct?

DUANE WESSELS: Yes.

BRAD VERD: We're in Section 9 and you keep referring to Recommendation 2.

DUANE WESSELS: I'm scrolling to get there.

PAUL HOFFMAN: Brad, I think the answer to your question is no, not yes. I think we're talking about the thing that is labeled Recommendation 2 which is at the bottom of that page.

DUANE WESSELS: It's at the bottom of page 25. Can you go to the bottom of page 25, Ozan? I may have confused things by adding some numbering earlier this morning.

PAUL HOFFMAN: You did, I put a note that about that, about five minutes ago.

---

BRAD VERD: We're taking Recommendation 2 and moving it to Section 9, is that correct? Is that what I just heard?

DUANE WESSELS: That's the proposal, yeah.

PAUL HOFFMAN: Not the recommendation that is labeled 7.2 by Duane this morning.

BRAD VERD: I have a suggestion but I don't have to throw a grenade in the room too much. This is good work and I'm not changing the work. I think that maybe we need to kind of follow what we did with 37 and 38, which is have a metrics document and then a separate document that is the recommendations, the set of recommendations that we're going to give the Board, if that makes sense.

Because this document, if we go back and reevaluate it, you don't want to update it and resubmit it with another and have the recommendations in it, much like we did with 37 and 38. 37 was the content, the meat, which is the metrics in this case and then a separate document would be the recommendation that we hand to the Board.

---

**RUSS MUNDY:** The current wording in the document, the recommendations are not directed to the Board, they are directed to RSSAC. The current wording in the document RSSAC is making recommendations to themselves and I don't know that we have identified any recommendations for the Board at this point and perhaps that's an item that the work party should undertake.

**PAUL HOFFMAN:** Sorry Russ, I disagree with what you just said. Reading the words, can you scroll down or up a bit, other way, stop. Recommendation 1 is a recommendation that the official implementation of the system must act, that's not a recommendation to RSSAC, that is a recommendation to future PMF or whatever jumble of consents we are considering. Recommendation 2 is also to that group, it's not a recommendation that they should do something, it's an acknowledgement to them that we're going to do something. If we want to separate out recommendations to us, it's really just recommendation 2.

Recommendation 1 is -- I'm sorry? And 3, yeah, so 2 and 3 are to us. 1 is to a future jumble of consonants. I could see separating that out as Brad suggested, I can also see leaving it in here, just making it clear that what we are doing in recommendation 1 with the bullets are in fact summarizing what we have said above. It's really not a recommendation as much as a summary.

---

BRAD VERD: For clarification, who's going to go build this system? You're saying RSSAC is or we don't know who is yet?

FRED BAKER: I would presume that it's somebody designated by the GWG.

BRAD VERD: Why do we have to wait for 37 and 38?

FRED BAKER: I don't know that we do.

BRAD VERD: This doesn't have to wait for 37 and 38 right, this could be Board, we want metric on the root system, we've talked about it, we know what we want to collect but in order to collect it, somebody needs to go build the system.

DUANE WESSELS: Nothing in the work that we've done talks about who's going to build this, we're describing it such that it can be built.

BRAD VERD: I agree but the conversation here, that I keep hearing is, this is for PMMF, which means it's tied to 37 and 38, and I'm asking why?

---

JEFF OSBORN: I was just going to say the same thing. I'm wondering whether we didn't start the metric conversation because we had some things we were going to have to measure because of or is this just, hey measuring out of the blue is a good idea?

BRAD VERD: Again, I feel we need operational experience with these metrics. We have an idea in our head but we don't know what we're going to see when start measuring it. I feel like again, if we wait for 37 and 38 to then do some big bang approach, saying we had this document that we wrote two and a half years ago, let's go figure this out, I feel like we're already kind of failed.

When we started this, this conversation metrics, there was a lot of confusion, was it tied to 37, was not, was it, was it that? We had all agreed, I don't remember where we were, I think it was Barcelona maybe, where we all said, we should start this work in lieu of 37 and 38, like assume it doesn't happen, we still should define what metrics look like. I think if you look at the scope of the document, it says that this should replace 001 or be new document. Again, none of that's tied to 37 and 38, so I feel that there is some -- I'm not feeling the connection here and I just like to understand where it is.

JEFF OSBORN: If it isn't tied to 37 and 8, who's buying the vantage points?

---

BRAD VERD: Again, that could be RSSAC saying, look we want measurements, we need measurements to understand the health and the availability and what not, so ICANN Board, our recommendation is to build the system or contract with somebody to build this system and here are the requirements.

JEFF OSBORN: Independent of 37 and 38, this is just, go ask for money to buy 29 vantage points?

BRAD VERD: I'm not asking for money to go buy it, I'm just say, you go figure it out, this is -- what you've defined here is the requirements for what's needed.

DUANE WESSELS: Wes has been waiting patiently.

WES HARDAKER: Thank you. Bringing it back up a level and trying to summarize everything I heard in last 10 minutes? I've lost track. A couple things, one there's multiple recommendations in here and it's unclear which ones are going to who, so my recommendation is to label each one, recommendation 2X where X is a defined entity either now or possibly some entity in the future and if we know the GWG likely to be stood,

---

we could certainly put that in there. It could be the PMMF or whatever it may be in the future but somehow direct -- state who it's to and when if possible.

The questions are what actions are going to result of these recommendations? Most recommendations should have an ability to do something based on it, it's not just a piece of advice, it's I want you to do this. I think what I'm hearing from Brad is, do we want to ask the Board now for someone now, possibly the ICANN Org through quite likely OCTO to go set up a preliminary measurement system so we can start collecting data? If I summarized everything, I suspect most people agree with I just said.

RUSS MUNDY:

Wes, I think you get a winner there. Personally, as one of your co-chairs, I have no problem at all with that type of restructuring, I think we also agree that the recommendation 2 goes down into the future work section, so that really leaves what the current Recommendation 3, which says, RSSAC should relook at the document, the contents at some undefined periodic interval.

There's one recommendation for RSSAC and then the rest I think are all related to the go build a system, at least in my mind, that's something that our work in the caucus is going to RSSAC, RSSAC can make and normally makes their recommendations to the Board. If it's agreeable with folks, we should try to word it so that it is a

---

recommendation for the Board to examine how to do to this and do it promptly.

However they figure out to do it, they might say, this is got to wait till GWG is done but I think that if I'm reading Brad and Wes correctly, there's a fairly strong need or view on the part of the operators, the SO operators today that says, we want it sooner rather than later. Stating in that manner would probably be good. Is that a good consensus, Jeff?

JEFF OSBORN:

I feel like I've got my Simon Legry hat on here because I'm always the guy bringing up the money. I'm wondering though whether this isn't giving them a free cow when we're trying to sell the milk. At some point if we provide metrics, do they then say, never mind paying you, now we're getting all those responsibilities things we were looking for.

It's not as complete as what the SLA's were imagined to be I think but to the degree that we go in and start providing metrics and measuring things and all the rest of it, it's a degree of accountability we didn't previously provide and I thought we were envisioning this as a quid pro quo.

WES HARDAKER:

No, because there's no stick half of the metric. If some of us fail there's still nothing to hold us accountable. The whole point of 37 is to

---

build the economy framework so that the metrics and the accountability could go together.

JEFF OSBORN: I'm worried it's more like a slippery slope rather than it's a complete capitulation. If nobody else is concerned, I'm absolutely willing to be out voted. It nags at me a little.

BRAD VERD: I'm not trying to out vote anybody; I'm trying to understand. If I'm ICANN I can go build the system today, what are we giving them now? This is just a monitoring system collecting data. RIPE does a lot of this. I'm just curious, what are you capitulating?

DUANE WESSELS: The thing that's new would be the documented thresholds.

JEFF OSBORN: We're implicitly measuring ourselves and we've implicitly saying what is and isn't good to use the Brad phrase and so while there isn't a stick in their hand, we're setting up a target, we're painting circles on our foreheads, we're lining up by order of strength, we're doing an awful lot, so all they need is a stick.

Previously this didn't exist. We were pretty independent and capable of doing what we wanted to and we're sort of not really answerable to someone and this is, now we're getting report cards and now we're

---

being told what failure looks like, we're having thresholds. It looks to me like a number of steps toward something that we thought we were going to get paid to do.

BRAD VERD: Okay, then I feel like this conversation should have happened long before we even started this metrics work then. I'm not sure what we're doing then.

JEFF OSBORN: I've been looking for the turn off to the part where we put a financial model together out of this from the first meeting. I've attended every single one of them. I keep waiting for what's the right place to do that, so I'm with you there.

FRED BAKER: I think this is an important conversation to happen but if you don't mind, I'd really like to not derail this meeting to have. I'd like to finish the metric work.

DUANE WESSELS: I think there are no remaining issues in the document but we have some work to do and given that we have some more time, we have about another 30 minutes in this session, right? I wonder if we should back and revisit the discussion about latency?

---

This would be RSO Response Latency, which is Section 5.2. Again, the discussion is between 250 and 400 or 250 and something else and how do we get consensus on this?

BRAD VERD:

Can I just make one comment before we move on? Sorry to do this, I just -- it's in my head, I feel like I should say it. I feel that 37 we said as a group was the beginning of a discussion and a discussion that we wanted to drive which was why we wrote it and we didn't want to wait for somebody else to write it for us and hand it to us.

I feel this is the same thing, if we don't want to provide what we think the metrics should be and we want wait for the GWG to go figure it out, okay but that essentially somebody giving us what to do, verses us driving the conversation. I think there's a fine line here between -- that we need to walk, to figure out what is right for this system and what is best for everybody here. I'm not sure that -- anyways, I just wanted to point that out. This drives the conversation verses having somebody give it to us.

I will remind everybody, that we will be in the GWG, we will have a voice in there, I think we have a final say that comes out of there, so it isn't like they go off in a box and come back and all of a sudden something different and we don't know about it. We're involved, we're engaged and I feel that people have lost sight of that also.

---

WES HARDAKER: Can I add one more point? You said exactly what I wanted to say a minute ago and didn't, so thank you Brad, that was really well done. I'm going to add one more point, which is that we have decided, with I think very good reason and purpose that we are becoming a transparent organization, all of our meetings are open and so we are now developing more pieces to that whole system in the open so everybody sees or motivations and purpose behind it. It's not done. This is another piece to a not done system because we want to build as much of it as possible.

DUANE WESSLES: I guess the latency metric being one of the more contentious things that you could be measured on, we find ourselves in this situation where we don't have consensus on what the threshold should be. We have, I wouldn't even call it rough consensus, sometimes we have consensus and sometimes we don't, it's sort of flips back and forth. It's intermittent consensus, that's maybe a good way to put it.

As the work party chair, I feel a little bit -- there are certain things where I think Russ and I have used our discretion to move things along but this is not one of them where I just want to push it through without everyone feeling comfortable, that they've had their say and gotten their point through.

WES HARDAKER: With my sheppard hat on, I think you've guys have done great job meandering through most of the contentious issues and I know hasn't

---

been an easy working group to lead. I liked Russ's suggestion earlier of taking it to a large extent, to a mailing list saying, here's the current text for that section, we want to hear any objections people have and it would certainly nice if some people with a vote would say, no, no we're actually going to descent from that, hopefully that's not going to happen, we've actually had a great track record in the last few years of pretty much unanimous votes on everything.

I think one other thing that could be helpful in that note to the caucus because I think some of this is hard to read and understand, when they get down to 250, that's a hard cutoff on the upper end, if you gave some example scenarios of what would be good and what would be bad, here's a set of measurements and have a bunch at two seconds, yet 51 percent is under 250, so people get that median concept down.

There's a lot of math in a lot of these sections, which is all fantastic and I love reading equations but you can get that wrong and examples go a long way toward helping people understand exactly what this means. I think a lot of people are probably missing it.

DUANE WESSELS:

I like that idea too. I like including examples and graphs to clarify those things. We'd also hoped and actually still plan to have this referenced dataset that people could go and look at and sort of convince themselves that the metrics are being calculated in an appropriate way and that we all agree on those formulas and things

---

like that. Unfortunately, that dataset doesn't exist yet but I hope it will still. I like that.

I know Paul looks like he doesn't want me to call on him but you've been running proof of concept on this for at least a few weeks now, you've got a stable set.

PAUL HOFFMAN:

But I have not actually been -- I've been collecting the measurements but I have not been doing any metrics. I've been fitting with the correctness as well as the rest of stuff in my day job. I do have a bunch of data that someone else could use or sometime now between this meeting and Singapore, which doesn't seem really bloody likely but I could crunch through some of that.

We have the problem though -- I'm sorry, we don't have a problem, a wrinkle in this is, we also decided that these were going to be aggregated by month. Before, when you put up some of your data and I put up some of mine, I was doing it by day.

Now again, by month should be better unless a root server operator is being consistently close to the threshold line. What we saw in the daily data was, that 250 was generous. It's particularly generous for the couple of root server operators who've already included cloud providers because they get basically ones out of those, especially with the way that I'm collecting now, which is from large data centers, where the cloud providers also have systems.

---

If today somebody said, do I believe everyone would make 250 on a monthly basis, I would say yes from looking at the data. I can't commit to having stuff crunched through, especially since the meeting is going to be at the beginning of Singapore but I would guess that and if somebody has a lot of time right now, I am happy to send them the couple of gigabytes of data that I have -- of collected measurements that I have.

But again, I don't think that's how we're picking 250 or 400, I thought we were picking thresholds based on what we thought the service -- or a root server operator should be offering, not what's going to make it so that everyone's okay?

DUANE WESSLES:

Yes, I agree with that. I think we've mostly been able to do that. We have lots of different opinions and so getting consensus around all those opinions is sort of the struggle, I think.

MICHAEL CASADEVALL:

It's hard not to think -- compare what are my numbers and it's hard. One of the things that helped me was, what would we expect of a new operator? They may look at things differently.

DUANE WESSELS:

Yeah, exactly. I guess if there's no further discussion on this particular issue than we'll just have to take it to the list. I'm sorry, Michael, I didn't see your hand.

---

MICHAEL CASADEVALL: Just two points. Maybe there's an easy way to push this past this issue is, put whatever number we decide as tentative and then put an asterisk and say we'll discuss it once we have more data available? Paul, talk to me at the break, I probably have the time to crunch through that data.

DUANE WESSELS: We've still got a few minutes left, was there anything else that we needed to talk about, Russ or Steve, that you can think of? The caucus meeting is when?

STEVE SHENG: It's on Tuesday, in the afternoon, let me tell you the exact time.

DUANE WESSELS: I don't need the exact time. I guess that gives us enough time to update the document with what we've talked about today and present it to the caucus on Tuesday, I think that's reasonable. I don't think we can necessarily come up with examples by Tuesday but maybe the end of the week or next week, something like that.

STEVE SHENG: Thanks, Duane. I think perhaps we have smaller meet to finalize the changes that are agreed today. I think the discussion towards the end, about how to structure the recommendation is an important one that

---

we need to think a bit more. I think the hard work of specifying it but how to structure the recommendation section is very important, so we probably need to give a little bit more thought on that area as well. Thanks.

RUSS MUNDY:

I know we wanted to move on past this but the sense that I got from the work party is that, it is generally agreed that what's currently Recommendation 1, would be shaped to go to the ICANN Board with the general message being, sooner rather than later is better, without saying the specific words, rather than tying anything to the GWG or 37, 38 process, am I interpreting what I heard correctly? I know there was some concerns about that. I'm sure Brad would agree. Anybody disagree with structuring one for the Board? Paul?

PAUL HOFFMAN:

I'm not disagreeing, since I may be on the pointy end of that stick, I would like a little bit more definition, for example, is this you want the Board to initiate this collection and continue it forever? Do you want them to do it for a year for examples? Do you want the Board to do this until the PMMF has started up? That would be useful to me but I also think it would be useful to the Board in order to say for example, figure out where this would fit into the budgeting process.

DUANE WESSELS:

Liman?

---

LARS-JOHAN LIMAN: I suggest we go for a limited term, maybe one or two years because at this point, we don't really know where we want to fit this in the hopefully forthcoming model and we don't really know either what the model is going to look exactly like and we don't know when it's going to happen. If start with a limited term, we can then ask either to extend it or change the model, change the process for collecting or so on. I propose a two-year term.

DUANE WESSELS: Wes?

WES HARDAKER: To quote somebody else at this table, the first thing they're going to want to know is, how much is going to cost? Either you have got to come up with a cost mechanism for that or you've got to say, we recommend you get a cost from OCTO, who's likely going to be the stuckie for doing it because they have to make a decision and they're not going to say yay or nay without knowing how much this is going to cost.

BRAD VERD: If you want to leave Section 7 alone, I volunteer to go draft a new Board Recommendation Sheet and share it with you guys separately if that would help. If not, then that's fine.

---

RUSS MUNDY: Oh yes, please, if you're willing to go take a whack at it, yeah, thank you.

DUANE WESSLES: Alright, I think we're done, unless there's any last comments, then we can break early. Ozan?

OZAN SAHIN: I have two updates before we break. One is that we have the How It Works Tutorial after this one. You're now in 517C, so How It Works Tutorial for Root Server Operators will be 517D starting at 3:15.

The other one is about the dinner, we are having the RSSAC dinner tonight at 7 in restaurant [inaudible]. If you want to walk to the restaurant with the group, we are going to meet at DoubleTree Hilton on sixth floor and depart by 6:45.

RUSS MUNDY: Okay, unless there's a last-minute thing, we give everybody 12 minutes back. Thanks.

**[END OF TRANSCRIPTION]**