OZAN SAHIN:	Good morning, good afternoon, and good evening, this is the RSS Metrics Work Party teleconference on the 19th of September, 2019 at 1700 UTC. On the call we have Duane Wessels, Jack Biesiadecki, Jeff Osborn, Kaunori Fujiwara, Ken Renard, Mauricio Vergara Ereche, Paul Hoffman, Ryan Stephenson, Shinta Sato, Dave Lawrence, and Matt Weinberg. From ICANN Staff we have myself, Ozan Sahin. I would like to remind you all to please state your names before speaking for transcription purposes, and thank you, over back to you, Duane.
DUANE WESSELS:	Thanks very much, Ozan. So, welcome everyone. The plan, things that I would like to accomplish today is to go through the document sort of briefly, spend maybe half our time on that and the other half on the spreadsheet where we asked Caucus members to fill in their ideas on thresholds. So, Ozan has put the document into the Zoom for screen sharing, hopefully everyone can see that. Ozan, why don't you scroll down a little bit, until we find something interesting.
	Alright, so, some of these things, I don't really want to spend on time today, but I will say that I would very much encourage all of work party participants to take a pass through the document and see where comments have been made, or maybe even read the uncommented sections, and if you see something that stands out, please make a comment. As we work through this with Staff, we're going to take any sections that don't have a lot of work party discussion, we'll take those as agreed to and start resolving some of these comments and accepting

the changes, and so on. So, please take a pass through it if you have time.

This first section is something that we're saving for discussion until the RSSAC workshop in a couple weeks. This one, as well, we're not really going to talk about that today. Let's scroll down a little further. So, here's a comment from Paul which is forward looking to the future about where vantage points might be located. I think this is fine, and again, if there is no discussion about this, then we'll accept that in an editing pass.

So, one note here, this section 4.7 which is about measurement reuse. Originally this was written in a way to give an implementation of these metrics some flexibility, that they could reuse measurements from one metric to another. Paul has been going through and proposing some changes that make that a little bit stronger to say that actually the measurements will be reused, so there is no need to make them twice. That's fine with me. So again, if there is no additional discussion at this point we'll probably move in that direction and clean up some of the comments here in the future.

PAUL HOFFMAN: This is Paul. So at this point if people are wondering like what does this refer to, at this point we've got in every five minute interval one check that handles both availability and response latency, and then one or two checks for correctness, depending on how we do it. So, I just want people to get that in their head, that if we agree that we can coalesce, we still only are having either two or three checks per five minutes. But

where the difference is, and as we get to it, I want people to think about we know that we're checking for UDP, PTCP, V4, V6 for that first check, root SOA, what do we want to do for the other ones, as well? Okay? Does that make sense?

DUANE WESSELS: Yeah, I think actually those SOA scores are going to be used for three metrics, right? It's availability, response latency, and also publication latency.

PAUL HOFFMAN: I'm sorry, correct, absolutely correct. So that's for those three, so we'll have four to do for those three, and then the question either we'll hit now or later will be, and how many, do we want to do one or four for the one or two correctness?

DUANE WESSELS: Right.

PAUL HOFFMAN: Okay, great.

DUANE WESSELS: Ozan, let's scroll down a little bit further. Alright, so we're getting into the section of the root server metrics, and these have been pretty stable in the recent meetings and calls. So I think this comments here is actually about what we just talked about, reusing the previous queries.

So, that's fine. There's a little edit from me, again, this is all related to the idea of reusing measurements from a previous metric, so that's kind of what this edit is about.

Okay, so we're getting into DNSSEC correctness here. There are some recent comments from Paul and myself about originally it said that the vantage point can choose which transport and we had a discussion about maybe doing something different there, because otherwise the vantage point could always choose UDP, or whatever. So the direction we're moving now is to say that the transport should be chosen at random, that's one option, so that you get some coverage. The other option is that we could say that these measurements must be done over all the transports. So you would do a DNS check over UDP V4, UDP V6, TCB V4, TCB V6 and if we do that then we have to decide do we aggregate all those together or would be publish correctness, like transport specific correctness measurements, I guess.

So, that would be the discussion there. Personally, I would advocate for doing a random selection, and again, if you have thoughts on that, you could raise your hand now or you could leave comments in the document. Alright, not seeing any hands, so we'll continue on.

There is a clarification about TLDs here which is a good change, so I think that will be accepted. Go ahead, Matt.

MATT WEINBERG: With regard to the transport mechanism, we said that we will be doing separate availability checks over multiple transport mechanisms, right? Did we say that?

DUANE WESSELS: Yes, and publishing different results over different transports, yes.

MATT WEINBERG: So, do we care about the transport mechanism for correctness? If we're already checking for availability, can we just settle on one transport for correctness check? I'm asking out loud here.

DUANE WESSELS: Well, I think the concern would be that if this document said that for example you always do correctness over V4 or UDP, and then if somebody wanted to be malicious, they could be malicious over the other transport and sort of get away with it.

MATT WEINBERG: Okay, okay, I get it. Thank you for the clarification.

DUANE WESSELS: That's the concern.

MATT WEINBERG: I grudgingly agree, okay.

DUANE WESSELS: Okay, let's continue on down a little bit more. Section 5.5.4 is root server correctness by matching and I think we talked quite a bit about

this on the last call and agreed to include this. You will notice that there is quite a lot of edits and changes here. The section that we're looking at now, I think was mostly written by Paul, maybe with some edits from me here, but I took a stab at sort of rewriting your text, Paul, a little bit more in the style of some of the other metrics, so that's below.

Scroll down, Ozan, please. It starts with something like DW suggestion below. I tried to capture the same spirit of what you were proposing and wrote in the style that's a little bit more consistent I think with the rest of the document. I don't know if you've had a chance to read it, Paul, or anyone else, but feel free to give comments on that.

PAUL HOFFMAN: I have not had a chance to read it, but I will, and I suspect it's just fine. What we need to come to agreement on are some of the things that would be in common between the DNSSEC correctness and the matching correctness. But they should alike, so that, so that, I mean, an implementer will be able to do whatever we write, but we want this to be readable by somebody who is figuring out do they agree with our metrics and our threshold. So, they should like alike.

DUANE WESSELS: Yeah, and so a couple more comments about this. Paul and I had some conversations in the chat and I think even some email about how this might work, and I think we've settled on the idea that the vantage points will obviously issue these queries and collect the responses, but then they sort of pass those on to some central processing system for doing the actual correctness determination. So, the vantage point itself

does not really need to know whether the response was correct. It doesn't need to have a copy of the authoritative zone data. That authoritative zone data can live on the central machine, and that can happen there. So, that's the way this is phrased now.

There was another point I was going to make about this, but I forgot it, so I guess we'll move on. One other thing that's still sort of undecided is, as Paul has pointed out, we could extend this matching base correctness most likely to have the same functionality or serve the same purpose as the DNSSEC correctness. So going forward, we need to decide if we want both or just one. Okay, let's continue on down, Ozan, please.

Alright, so publication latency is another one where there has been some recent changes. The text here doesn't show all the changes because it got very messy and I went ahead and just sort of accepted some of them. But the idea here again is that the measurements for publication latency come from reusing the SOA queries in the other metrics.

This is one of the three metrics that will use the same queries and responses. Calculating publication latency again would be done on this central processing system, that system would have knowledge of when the root zone maintainer publishes a new zone, so that's kind of when to start the clock, and then it receives all of the responses and can do the calculation for each vantage point and each root server, when it did see the new serial number. Since these measurements are done on a five minute interval, that's sort of the granularity of this metric. You would be able to say whether the zone was published by the root server

within the same five minute interval or the next five minute interval, or so on.

So, that means that's just the granularity of the latency that we're going to get out of this. There's this text that's in purple here, there's a little bit of wiggle room here in how the central system can actually determine the start of the root zone publication time, either by subscribing to notify messages or by examining all of the SOA serial responses from all the root server operators. Paul, your hand is up.

PAUL HOFFMAN: Yeah, so if people are concerned about the five minute interval, just to be clear, on the proof of concept system that I'm set up, so far, unless I'm reading my own reports wrong, there has not been a single, and this is over the course of at least five days, there has not been a single SOA change that has appeared in every RSO, and this again is with 8 vantage points. There has not been an SOA change that has actually been done successfully in a five minute range. So I think five minutes is actually fine for granularity. Often it's 10 minutes, sometimes it's as much as 20 minutes.

The other point I wanted to make, and actually I should be saying this to Duane individually, the RZM contract doesn't allow saying that they're going to send notifies to other people, so we would need to figure that out if we were doing this as in some system that is not a root server operator is supposed to get notifies.

DUANE WESSELS: Yeah, that's right, that's one of the reasons I left this a little bit vague here, because there may be sort of beneficial implementation of this and there may be other implementations of these metrics by other parties, and it may not be realistic to have all of them subscribing to notifies, and things like that.

PAUL HOFFMAN: And Ray Bellis says in the chat, how would the potential phase difference between notifies and measurements be accounted for? To answer Ray, from my proof of concept implementation, I think that would happen where as long as the threshold for publication is more than one 5-minute interval, I think it doesn't matter. That is, if the notify comes in just before a time click, and therefore no one has the new SOA in the first set of measurements, it doesn't matter. It would look funny, but no one is going to be seeing those "look funnies." It's only if the threshold would be that short, and I can't imagine it would be that short. Hopefully that answers what you were asking, Ray. He says, okay, thanks.

DUANE WESSELS: Alright, good. Okay, if there is no more discussion on this topic, then let's continue scrolling down.

So, there's a change here in purple, again, talking about, it's sort of pointing out that this metric is a little bit different than the others, because in almost all the other metrics, we have one measurement in each five minute interval, so that's 288 measurements per vantage point, per server, per day. But in this case it's clarifying that the

measurements happen only when the root zone serial number changes. So generally that happens twice a day and so the total number of measurements to be aggregated would be about two times the number of vantage points per day. Alright, let's continue down.

The next section is the RSS metrics. There is a comment that we're leaving for discussion at the workshop. Keep going down, oh, can you go back up to the availability one? So we'll spend a little bit of time talking about this now, and maybe also when we look at the spreadsheet. As I and my colleague, Matt, were filling out the spreadsheet today, we had an interesting discussion about this metric, about what does RSS availability mean, and that sort of thing. The way that this metric is currently defined, it's a little bit strange, because the aggregation is just a simple aggregation of all of the individual servers together.

For example, if there is consensus around the idea that, say, it's alright for one of the 12 root servers to be unavailable for the course of the day, right, RSSAC has a document that's along those lines, the unavailability of a single root server is not a super big deal, and if we took that and it actually happens and we aggregate the measurements from a day when that happened, then as written, the RSS availability for that day would be quite low, it would be 12 out of 13, which is something like 93% or so. So, to me, that's a strange way to think about availability of the system as a whole. It might be better to think about it in terms of even when we had 12 out of 13 available, that meant the system as a whole was available for that day.

I see that Ray has entered a comment in the chat which says RSS availability means that at any moment of the day, could any root server be reached. So, if I understand what you're saying, Ray, you're saying that if any one out of 13 or how many is available, then that means the system as a whole is available for that five minute interval. Go ahead Matt.

MATT WEINBERG: Here's my opinion for what it's worth, and people can agree or shoot it down. The root server system should always be available. The question is what do we define as available? I would argue that for example 12 out of 13 servers up means that the root server system is available. Of course, that presumes that recursive servers will retry, but just bear with me here. So building on what Ray said, and I agree with Ray, Ray said one is too few, I would agree, one is too few.

I think what we should really decide is what is the minimum number of root servers that should be available at a given time slice and if we hit that threshold, then the root server system is available. It should always be available, the question is how many root servers need to be up or available for us to be able to say the root server system is okay. That's my thought, I could speak more to it.

DUANE WESSELS: Thanks, Matt. Paul, your hand is up as well, so go ahead, please.

PAUL HOFFMAN: Okay, so I agree with Ray because as far as we know, a typical resolver that is relying on root server P, if root server P goes down, the resolver will go to another root server within five minutes, most likely it will go within a few seconds, or if that root server has just rebooted, and it does a prime, and it starts looking for its favored root server and it hits root server P and P is down, that resolving is not going to wait five minutes before starting something else.

So, to that resolver, in either of those two states, five minutes is just fine. So, that to me supports Ray's assertion that if anyone was up in five minutes, that means it's up. Hopefully it's more than one, because that one will be really busy, but to me that is a reasonable availability for the root server system.

DUANE WESSELS: Okay, Paul, so you think that one is an acceptable number if we were to continue down this path?

PAUL HOFFMAN: I do, although Ray just said one is perhaps too few. I'm trying to do this based on verifiable numbers. So of course I wouldn't want it to be just one, but it would be sufficient to say that the root server system is up for any given result.

DUANE WESSELS: Okay, so when Matt Weinberg and I were talking earlier today, he found this formula, right Matt? Some actual math you can use to say if I want X availability and I have N sites, then M of them have to be up to give

me that availability. So, I think using a formula like that, we can either work it in either direction. We could say, for example, well, if we have 13 today, or let's say 5 is our magic number. If we need 5 to be up or we want 8 to be up, or 10 to be up, whatever that number is, we could choose that number and from that number we can calculate what the availability should be.

Alternatively we could say we desire an availability of 9.999 or whatever, and if we have 13 RSOs, then it means N of them to be up, and the formula will tell us what N is. I see there are some more comments in the chat. Russ says if we decide that 1 is too small, then I think we also have to say how many vantage points have to see more than 1. Yeah, it gets a little bit complicated, since our measurements are on these five minute intervals, in one interval we can calculate, oh, I saw this many RSOs up, but then to me it's difficult to propagate that up to the course of a whole day, because is being 5 out of 13 in one 5-minute interval, how do you express that in terms of a day when you have all these other five minute intervals where maybe they were all available?

More comments from Ray, if one root server is reachable, it will get so much traffic that the latency measurement will fail its threshold. That's a possibility, I don't know how likely that is to happen, it's probably never happened, but yeah, I could see that happening. On the other hand, if one root server is handling all the traffic without increased latency, then arguably the availability count is moot. Matt says I think the true question is what is the minimum number root servers that should be up for a given time slice? You mean a five minute interval, Matt, right?

MATT WEINBERG: That's right. I think one is too little, but we're probably not going to solve it here on this call. Will this be discussed more at the workshop here in a couple weeks? Because I can show the math...

DUANE WESSELS: I think it needs to be.

MATT WEINBERG: Perhaps we can have a little 15 minute session about the math behind availability and what it would mean of a minimum number of root servers were up and we could assign a percentage up time per root server, then the RSS as a whole will be 99.999% up. And there is math behind it. I think that would be helpful for everybody to see.

DUANE WESSELS: I think so, too. Paul, is your hand up again?

PAUL HOFFMAN: Yes, it is. So, I'm having the problem in this discussion that I've had in previous discussions which is we have measurements, we have metrics, and then we have thresholds, and I think the last five minutes we've mixing up measurements and thresholds. So, I think the RSS availability measurement should be something that is directly measurable, not desirable, and we wait for discussion of thresholds to talk about desirability.

DUANE WESSELS: Okay, I think in the interest of time, Paul, you and I should maybe follow up separately later. It's a fair point, but I think it's got complications. So, we can talk more. I'm going to try to stay on time. So let's scroll down a little bit more, Ozan, let's see what else is down here. Okay, I just noticed that it looks like we need to fill in section 6.4, well, let me ask the group, we have a correctness by matching the metric for the root server operators, do we also want such a metric for the RSS? And if so, we need to write some text here. Any opinions? Well, if you think about it and want to add some comments in the document later, please do so.

I want to now skip down, so there are some edits to the recommendations, again, please read these. I'm not going to go through them at this time, but I want to point people in particular at the examples. So there is a Section 8 which is example results which has been totally rewritten from the previous versions. These examples are presented in the style of assuming that the only thing that gets published is whether or not threshold was met or not met.

So, looking at the first one as an example, the metric is RSO availability, there is a date, so this is for a single day. It lists under the RSO here, it lists N root and there are four rows for the four different transport combinations. Then there are these green and red boxes which says the first one IPV4 UDP, the measured performance or the metric was above the ridiculous threshold of 15%, so greater than or equal to 15%. For TCPV4 it was above the similarly ridiculous threshold of 12%. But for

IPV6, the metric was below 15%, so it's shown in red. And then again for the next one, it was above.

And then the next column shows the number of measurements that were included in that metric. So, this is for example 20 vantage points times 288 measurements per day for that root server. So is that sort of clear enough to everyone? Any questions or suggestions about how the examples are presented?

I will say that when I wrote these, I did try to choose really ridiculous example thresholds so we didn't get too focused on thinking that these might be actual threshold recommendations from the work party. Alright, you can scroll down a little bit Ozan and just show people that there should be an example for almost every metric, I think there are lots of them to look at. Alright, thank you. Would anyone like to discuss anything about the document at this time before we move on to the spreadsheet? Any concerns or wonderful suggestions to consider right now? Okay, I guess not. So, Ozan, let's switch over to the spreadsheet, then.

So, if you haven't had a chance to see the spreadsheet yet, this is what we've asked caucus members, any work party member to fill in some values here. Ozan, if you could quickly switch to the instructions tab and just show people that is there. So there is an instructions tab which is designed to guide you in how to fill this out. It explains a little bit about what the different rows are. Go back to the actual spreadsheet.

So you can see here in Column A, there are these two sections; there is minimum performance thresholds and good performance thresholds.

This gets back to the discussion we had on the previous work party call which is what sort of thresholds should we be considering.

In terms of the RSSAC037 work, that is targeting the minimum performance threshold idea, and that's where most of the people have provided answers. All the metrics from the document are listed here and you can enter your name and start entering values. If you'd like to make a comment about why you chose a particular value, or your justification, you can enter a comment into the cell. The fields with comments have these little yellow orange triangles, and if you hover over one of those cells, the comment will application.

When you go through this, do keep in mind that almost all of these metrics as they have been defined, the threshold refers to the median value of the distribution of all the measurements. So the latency one is an obvious one, and it says there, the threshold is on the median of all the latency measurements. For the publication latency, it's also median. I think everything, any time there is a distribution involved, we're proposing to use median values at this time. Any questions about the spreadsheet before we start to look at some of the actual answers?

Okay. Also, I'll quickly point out, in Column B there is a fake entry, which is designed to just be an example of how to fill it out. So John Q Caucus is not a real person, and the thresholds suggested by John Q Caucus are a little bit ridiculous, but it's just to show you what the format should be. So, you'll see that different people have provided different amounts of answers. Some people think that maybe certain thresholds should not be provided in this work party or it doesn't make sense, or it cannot be answered as asked, and so on. One thing to point

out and get out of the way is for the correctness metrics, pretty much everyone is saying 100%, there is not a lot of variation there. Steve, did you join the call?

STEVE SHENG: Yes, I'm on the call.

DUANE WESSELS: I just got a popup that said Steve is now the host, and then the sharing of the spreadsheet went away.

PAUL HOFFMAN: Ozan was sharing.

STEVE SHENG: Really? Okay, hmm, I didn't do anything.

DUANE WESSELS: All you did was log in.

STEVE SHENG: No, I've been on the call for three minutes.

PAUL HOFFMAN: Oh, I can see, Ozan fell off the call and therefore posting fell over to you, Steve.

DUANE WESSELS: Oh, that makes sense, okay, got it.

STEVE SHENG: Okay, alright. So let me pull it up.

DUANE WESSELS: I can share it as well, if you would like, Steve, should I just do it?

STEVE SHENG: Yeah, please do.

DUANE WESSELS: Alright, for correctness, we're seeing a lot of 100%, which is not too surprising. There is a little bit more variation in some of the availability cells, and a little bit of variation in the response latency cells. We'll sort of point out something that I didn't think of until I actually went to fill out the survey for myself, we should have added separate rows for UDP and TCP, it's pretty common to allow for different latencies there, because obviously TCP has this connection setup delay and the metrics specifically call that out.

So you'll see the ones that I filled out, essentially that the TCP is always twice the UDP threshold. So, if you haven't had a chance to fill this out, I would encourage you to do so. We're going to take this as input to the RSSAC workshop when we again try to have the threshold discussion there. I think it would be really great for all of the work party members

to put in their suggestions here in advance of that workshop. Paul, or anyone on the call, should we talk about why you made particular choices or should we just leave this as it is? I know Warren is not here. Ken, I think you're here. Does anyone want to speak to their particular suggestions at this time? Paul, your hand is up?

PAUL HOFFMAN: Yeah, so I sent a four or five paragraph message to the list last week after Steve had put up the link to this, because I'm hoping that we actually put a rationale with whatever numbers we come up with, that we actually come to them by a rationale and that we actually publish the rationale in the document. So, my numbers came after I wrote the rationale, I literally had done some numbers, stopped doing the numbers, wrote the rationale, and did it. And I'm not expecting everyone to agree with my rationale, in fact, I can see that some people don't, because some people's numbers would indicate different ways of looking at it.

But my feeling, or the summary of the rationale I said was that correctness was more important than publication latency, but publication latency was way more important than availability and response latency, for the reasons that we have talked about all along and I mentioned before, which is if for a resolver, if either the availability or the latency sucks, they're just going to move on.

So that's where I came to where the availability to me, 90% or 95% is just fine, that's good enough. We don't want anyone who is only available 5% of the time, but if there are, and again, we also have to

deal with RSSAC 042 which allows a root server operator to have whatever architecture they want. They may have an architecture that is flaky for a particular design goal. We don't want it to be horribly flaky, but 90-95% is just fine.

Going to your question, Duane, about response latency and having a difference between UDP and TCP, I actually think that the setup time for TCP under normal circumstances is going to be under 50 milliseconds, so that gets sort of lost in the noise for my suggested threshold of 1 second, and my suggested threshold of 1 second, which some people agreed with, other people thought was too long, I can see from this, is I think 500 milliseconds, or half a second is enough time to get from any vantage point to the furthest, through the fat pipes of the network to the other side of the world.

And at that point, we again, gallbladder RSSAC042, which allows somebody to have any architecture, if there are thin pipes between, let's say that a new RSO is set up, let's call it Q, and it is really specific for they're going to have a whole bunch of instances but in one country, that's their goal, and they have let's say two major networks interchange points, but between those interchange points and where they put the root servers are crappy connections, that could be another half second, so adding those two together is how I got to 1 second. Hopefully that helps.

DUANE WESSELS: Okay, thank you. I saw a comment from Ray in the chat, still thinking about the publication latency and he's noticing that a lot of spreadsheet

were saying that one hour is sort of the minimum performance threshold which is very different than the root zone expiry timer which is 7 days. So that's quite a gap. So how do we explain this difference? Is the root zone expiry timer wrong? And should it be closer to an hour? I guess that's the question you're asking, Ray. Or should the publication latency allow for much longer time periods? Paul?

PAUL HOFFMAN: I tried to make all my thresholds sort of match a 95% expectation. The shortest TTLs in the root zone are 24 hours, let's call that 25 hours, and so I wanted to have the root zone publication latency or somebody who is missing that latency, affect no more than 5% of the users, so that 95% of the people will still have whatever change in a root zone cached. So I picked 1 hour as approximately 95% of the shortest TTL in the root zone.

DUANE WESSELS: Thanks, that's interesting. I came to 1 hour a little bit differently. I looked at the SOA refresh time, I think it is, which is 30 minutes, so I came up with 1 hour saying you shouldn't miss more than 1 SOA refresh interval, right? If you miss 2, then that's bad. Well, there's lots of stuff in the chat here. I'm going to follow up with Ray's discussion and then I'll get back to Ken. So, Ray said whichever one changes there, they're not currently what I would consider in proportion with each other, and I think that's a fair point. Mauricio says he reads the SOA on the publication latency as the fatal mark, the expire value as the fatal mark,

while the minimum being defined here is more like a warning. Thanks, Mauricio.

Scrolling back up on the chat, Ken, you said the main comment on the threshold spreadsheet is for RSS metrics, which were captured in the discussion a few minutes ago. So you feel like that's already been covered, then, I guess, Ken. Okay, Ken says yes. So, getting back to something Paul said, I think that's a very good point, when we do get to the point where the document has specific thresholds to recommend, we really do need some justification text around those. And so maybe this spreadsheet will be the starting point for those justifications, that would be great if people could fill in comments on the cells, that would probably be helpful when it comes time to putting these into the document.

Question from Shinta, when is the next RSSAC workshop? It's actually in about two weeks. We'll be meeting in Reston with the RSSAC members. Almost all the time at the workshop will be dedicated to this work. I believe that's still correct, right Steve and Russ? We have at least two whole days to talk nothing but metrics. So the work party members will be invited to participate in those calls, I expect most people to participate remotely, if work party members are able to get themselves to Reston, they're welcome to participate in person, but there is no travel support provided. So, between now and then, please take a look at that document and the spreadsheet and fill in as much as you can.

I think Russ said he is in place where it's hard to talk, it's noisy or hard to talk. Russ, is there anything that you wanted to add to the meeting before we wrap it up here? Also, I know Russ was maybe suffering from

a little bit travel fatigue, just coming back from a meeting on the west coast, I think.

STEVE SHENG: I think the meeting on the West Coast, the SSAC workshop is still ongoing.

DUANE WESSELS: Oh, it's still ongoing, okay, I didn't know that.

PAUL HOFFMAN: Duane, I just put something in the chat but I'll say it out loud, because it sounds like you're closing the meeting. We still have not defined at all what 'good' means. Not only define what we think the good metrics are, but we don't have any rationale for what is good. So, I'm quite concerned about the second half of the spreadsheet, because we don't have a basis for it.

DUANE WESSELS: Yeah, it's a little bit nebulous, it's a little bit hard to define. Do you think it's something you could take a stab at?

PAUL HOFFMAN: No, because I actually, I don't see it as part of our work. I saw that the work for 037 the thresholds for either getting thrown out of the root server system or if you're coming in, you have to show that you would be able to meet thresholds on day one. I didn't get the feeling from

anything in the statement of work that we were also supposed to be saying what was good for some value of good. And certainly in the last face-to-face meeting at the ITF meeting in Montreal, people had very radically different views of good, and I'm not going to speak for anybody else, but I heard in the course of less than five minutes, three very different definitions of good. So, us putting numbers here to me is a little bit scary.

DUANE WESSELS: Okay, I take your point. I guess that means that I need to come up with a definition of good and put it before the work party, because I doubt if anyone else would want to do that. Russ, in the chat you wrote that you agree with Paul that the work party is not required to define good, you mean define the thresholds or define anything about good? I guess, if nothing else, Russ and Steve, we should maybe revisit our proposed schedule for the meeting in a couple weeks and dedicate some time to this issue and get clarification from everyone about the extent to which we need to include good thresholds in the work party document.

Okay, so I think I've sort of run out of things to talk about. If anyone else wants to raise a point, feel free to do so now, raise your hand, or say in the chat, otherwise we will wrap up the call. Alright, thank you everyone for participating today, and we will see you in a couple weeks.

[END OF TRANSCRIPTION]