MARRAKECH – RSSAC Work Session: RSS Metrics (part 2 of 2)
Wednesday, June 26, 2019 – 10:30 to 12:00 WET
ICANN65 | Marrakech, Morocco

[DUANE WESEELS]: Okay, everyone, welcome back. This is the second session for the RSS Metrics Work Party. We're going to start off with another quick attendance since it has changed a little bit in the room. I'll ask to start again down here with Ken. You can just go quickly.

KEN RENARD: Ken Renard, ARL.

KARL REUSS: Karl Reuss, University of Maryland.

DANIEL MIGAULT: Daniel Migault, Internet Architecture Board, [DSO].

DUANE WESSELS: Duane Wessels, work party co-leader.

RUSS MUNDY: Russ Mundy, work party co-leader.

JEFF OSBORN: Jeff Osborn, IFC.

ABDULKARIM OLOYEDE: Abdulkarim Oloyede, work party member.

FRED BAKER: Fred Baker, ISC and RSSAC Chair.

CARLOS REYES: Carlos Reyes, RSSAC support staff.

OZAN SAHIN: Ozan Sahin, RSSAC support staff and also the Internet remote participation management. I see [Anand Ataunarula] is in the room. Also we have Kazunori Fujiwara, Keith Bluestein, Paul Hoffman, and Terry Manderson.

HIRO HOTTA: Hiro Hotta from WIDE and Paris.

LARS-JOHAN LIMAN: Liman, Netnod.

PAUL MUCHENE: Paul Muchene, RSSAC Caucus member.

RYAN STEPHENSON: Ryan Stephenson, DISA.

WES HARDAKER: Wes Hardaker, USC-ISI.

AKINORI MAEMURA: Akinori Maemura from the ICANN Board, observing.

UNIDENTIFIED MALE: [inaudible] from .[inaudible]

UNIDENTIFIED MALE: [inaudible], RSSAC Caucus member.

UNIDENTIFIED MALE: [inaudible], domain [inaudible]

SHINTO SHATO: Shinto Shato, JPRS, RSSAC Caucus.

RAO NAVEED BIN RAIS: Naveed, a member of the Caucus.

UNIDENTIFIED MALE: Hello, everyone. I come from China [Academic] of Information and Communication. Thank you.

FRED BAKER: Let me repeat a comment that I made in the first session. That is that this is a Caucus member. If you're a Caucus member, you're perfectly welcome to sit at the table with the rest of us.

DUANE WESSELS: Thanks, Fred. What Russ and I would like to do next is try to get some closure or consensus on the discussion we just had. The idea that I have is I'm going to pose two questions to everyone in the room. If you're in RSSAC in the Caucus, we would very much welcome your answer. We won't require you to answer. I would like to require you to answer, but I won't. But please answer.

Here's the two questions. Not surprisingly, they're going to be about probe locations. In your answer, you do not need to use the words "near" and "far." You can use whatever words you want. The question is, where should correct [mis]measurements be made from, and where should latency measurements be made from.

We'll start down this end of the room.

KEN RENARD: I think that latency measurements should be made from everyone, afar and near. For the correctness measurements, we may need to throw away "far" and collect, keep the probes all the same. But we may need to throw away far.

UNIDENTIFIED MALE: Correct measurements as close the server as reasonably possible. For other probes, geographically distributed somewhat evenly across the globe, located in more network centers than far, tucked away corners of the globe.

UNIDENTIFIED MALE: Latency I think should be made with a subset of the probes that are spread in the scope of the RSO. For example, the RSO is targeting a specific region. Probes should be in that specific region and so on. For correctness, I don't know because, if it's local and self-reported, it's very easy to cheat about the correctness.

DUANE WESSELS: Don't assume it's self-reported. Just where they should be made from.

UNIDENTIFIED MALE: Okay. It should be made at the point where no one is in between.

UNIDENTIFIED MALE: This is too interesting discussions, one theoretical and one practical. It seems like we're going to knuckle down and choose some preexisting probe network to use. So using this absent that framework seems kind of silly but interesting. So—

UNIDENTIFIED MALE: I don't think that's necessarily true, that we're going to choose some existing [inaudible].

UNIDENTIFIED MALE: Okay. Well, if we're doing it completely from scratch, then I think we should rank the countries of the world by population and then put this in the most populous cities that don't have a root instance until we run out of cities with a million people. The problem—

DUANE WESSELS: Is that for latency or for correctness?

UNIDENTIFIED MALE: That's for latency. For correctness, you can do it close, I think. If we're designing the diversity out of the system by having everybody be in ten big cities, that's silly.

DUANE WESSELS: I'm just asking you where these measures should be made from.

UNIDENTIFIED MALE: The middle of nowhere.

DUANE WESSELS: Okay.

RUSS MUNDY: We'd like have everybody's input, so would you please give us your view on these two questions?

ABDULKARIM OLOYEDE: Okay. I think I share his views. Probably I think measuring – well, my fear is that measuring from the middle of nowhere would actually mean, from where is the middle of nowhere?

RUSS MUNDY: For both correctness and latency?

ABDULKARIM OLOYEDE: Yes.

RUSS MUNDY: Okay. Thank you.

FRED BAKER: One bias that I have, one viewpoint that I have, is that it should be possible to use these metrics from anywhere. I mentioned the possibility of do I need a root server and being able to conduct the test and determine whether you would be in a better case from doing that.

In the correctness thing, I might find that an RSO is normally delivering correct information, but somewhere in the great wide world there is some bit of middleware that is corrupting that information in some way for whatever reason. So being able to prove the case becomes interesting.

So I would argue that anything that is an SLA measurement of an RSO needs to take into account the part of the network that the RSO is in fact in control of and can do something about. And it should be in that area. So latency wants to be from there, and correctness wants to be from there.

When it comes to the user experience, though, I don't think you can measure it in the same rack as the root server. I think you're going to need to measure it somewhere else. I almost don't care, but somewhere else.

DUANE WESSELS: Remotes? Yeah.

OZAN SAHIN: Paul Hoffman's hand is up in the Zoom room. Please go ahead, Paul.

PAUL HOFFMAN: For latency – Duane, I know you're going to have this – I think the question is relevant, again, due to RSO selection. So I would say, for latency, the probe should be self-selected by the root server operator.

For correctness, I'm also going to have probably an unpopular opinion here. I think that the correctness probes should be "far away," and there should not be thresholds on them. I think they should be strictly informational.

OZA SAHIN: Terry is up next. Terry?

TERRY MANDERSON: Thank you. From a pragmatic point of view, I like the idea of having probes for latency, having probes very near to root servers, perhaps even in the same rack as the root server, but not measuring that root server. So measuring every other root server. And that's the near aspect that I like.

I would also advocate having measurements of latency from as far and as wide as you possibly can because I want to know what the worst-case scenario is.

In terms of accuracy, to be brutally honest, I don't think it actually matters. Thank you.

DUANE WESSELS: You said accuracy, Terry, but I think you mean correctness?

TERRY MANDERSON: Yes, I'm sorry. I did.

DUANE WESSELS: Okay, thank you.

TERRY MANDERSON: Thank you.

UNIDENTIFIED MALE: For correctness, I think it should be a reasonably near place to the servers. For the latency, a various number of probes should be done for that from perhaps reasonably large cities or something like that. That'd be a good place.

HIRO HOTTA: For the correctness, from nearby probes. For the latency, measured from a subset of the probes. But I'd rather think root servers may be categorized or self-declared into one of the classes, such as [first] categories – for example, RSOs with a small number of instance that are focused on a specific region may say, "Assist me with 35% percentile [inaudible." Or the [group of the diverse] RSOs may say 100% percentile. So each RSOs should [inaudible] serve/select the category and such level is assessed by that. That's my idea.

LARS-JOHAN LIMAN: I'm very much with Carl here. I that correctness should be measured near and latency from a large number of probes spread over the global – with cutting of the trailing edges from that.

It struck me as we were speaking here. There is a Swedish idiom that says you shouldn't curse in church, meaning you shouldn't say the inappropriate thing in the wrong moment, but what about implementing these measurements? Ask the resolver implementers to measure this in the resolvers and have them report that back. Somehow – of course being an optional thing that we can ask people to please turn on reporting into this system because that's the actual point we want to measure from. Thank you.

RAO NAVEED BIN RAIS: I think latency should be distributed across all the world, but it depends on how we define near and far. That is what I'm still confused about. What do we say is near and what do we say is far. But it should be collected from probes distributed all over, just to have as close to accurate idea as possibly about the latency.

For the correctness, we would like to validate or to know that the root server is serving correctly as not as soon as possible. So that should be closer to the root server. So it might [inaudible], and that should be near.

DUANE WESSELS: Thank you.

UNIDENTIFIED MALE: I want both accuracy and latency to be measured by a geographically-distributed perspective with [inaudible] another level of geographic diversity.

PAUL MECHENE: I'm just going to give my brief opinion. For latency, as most have mentioned, I'll also agree with consensus: it should be measured from anywhere, where possible. For correctness, much closer to the root servers.

RYAN STEPHENSON: Hi. Is correctness even needed with DNSSEC? With correctness, does that mean that the root server operators or root server system or the RSSAC does not trust the DNSSEC technology? That's the question that came to mind with correctness of somebody serving whether these delegations or whether this record, depending upon whatever it is [that's queries], is correct.

But, if I had to pick near or far whatsoever? Within a decent geographical region with the understanding that those probes may not be of the same ISP and that there still could be some type of a middleware change or something. Maybe those resolvers like to modify something in the answer back to the clients.

With that being said, again, I guess I would be more self-reporting upon the RSO for correctness. For latency, anywhere with caveats. Again, there's a lot of things that root server operators can't control, so that of course would have to be noted.

WES HARDAKER: Where should correctness be measured from? Measurement of correctness – I thought your ideas there were interesting – of an RSO should be near because you're trying to correct their correctness. Measurements of the system, on the other hand, should be far to determine whether correctness if being properly deployed or interrupted throughout the world.

Latency of an individual root to determine if they're performing or if their infrastructure is up to par near their system in order to test that system. Whether or not they're geographically diverse is an independent goal. If any RSO's primary proposal of standing up their service is to achieve worldwide small latency, then they should be tested at a distance. Latency measurements of this system as a whole should be done from random locations to test whether the system as a whole is meeting the goals of the system as we want it to be.

DUANE WESSELS: Thank you. Ozan?

OZAN SAHIN: We have a comment from Kazunori Fujiwara in the Zoom room, which I'll be reading now. He doesn't have a mic available. "My idea is to measure from anywhere. However, if you need [latency SLA], compare latency for the root and latency for other servers. For example, .net servers or ICANN servers."

DUANE WESSELS: Okay, thank you. Does anyone in the seats want to give an answer?

No? Thank you for those. I think we've got an hour-ish or so left, right, Russ? There's—

OZAN SAHIN: [inaudible]

DUANE WESSELS: Oh. Go ahead, Ozan.

OZAN SAHIN: We have one more comment in the Zoom room from Keith Bluestein. I'm reading it out loud. "In case we don't call for the phone, I agree with Liman's approach. Latency should be measured all over. For correctness, I think there's benefit from monitoring all over to identify other integrity issues."

DUANE WESSELS: Thank you, everyone. I think we're going to close that topic on probe locations for now and move on to some other things. One of the other things that we want to bring to the attention of the Caucus and everyone here is that, in the current document, in the section that talks about metrics for the RSS, in a number of those, there are alternative approaches. One of the approaches is generally to take the measurements against the individual

services and then aggregate them appropriately using math and whatnot and coming up with a metric for the RSS.

The other alternative approach is to send queries to a local recursive nameserver and let the recursive nameserver do what it's designed to do and pick a root server to query and things like that.

Russ and I were talking to ourselves yesterday, and we talked ourselves out of this idea of the local recursive nameserver. I think it's got a lot of complications. One is you have to pick an implementation or implementations – BIND, Unbound, etc. – to put on here. Then in a sense you're just measuring how that implementation works. You're measuring its characteristics.

Another thing we realized is that doing it this way might make the measurements not repeatable and consistent over time because, as probe software gets upgraded, you might switch to a new version of the resolver and then it may differently than it behaved before, so you can't compare the past to the future and so on.

So we are proposing to eliminate those as alternatives in the approaches, and we'd like to open it up for discussion. If anybody would like to argue for keeping them, we'd very much like to hear that. I don't think we would necessarily remove these right away. We would take this discussion to the list as well. But now would be a good time for people to have an opinion on this.

KEN RENARD: Our objective here is really to measure the RSOs and RSS. I agree that that's the right approach, not using a resolver. The user experience question is a different question. That's probably better answered through resolvers. But I think we're on the right track, given the scope.

JEFF OSBORN: I hope this isn't widely off-topic, but if we're rethinking what is it we're trying to accomplish, I was going back to why are we discussing [inaudible] SLAs in the first place? What we're all talking about is, what is the state of our devices in a normal state?

When this all started, it was because , "Oh, hell. There are multi terabit attacks and the root service might be taken off the air." I don't hear any of us doing stress testing of any sort that would take something off the air and see how we react to it. I'm just being concerned that, when we go through months of doing this and say we've got it all figured out and they say, "Terrific. What happens under a big attack?" and we say, "We don't know. We have an ambient steady-state health monitor" … So is that anywhere in scope of this project?

RUSS MUNDY: Well I don't think that this is within the charter of the work party. Once a set of metrics gets defined and some values get identified as the norm and the thresholds and so forth, there'll be more information available to examine in the case of these large attacks. But I don't think it's anywhere intended to do anything to prevent or even necessarily monitor the large attacks. I think it's a related but separate question, I think, that you're raising.

There are two RSSAC shepherds of this, Brad and Wes. Do you agree that was how the work party was chartered?

FRED BAKER: I agree that that's how the work party was chartered. Frankly, I'm not sure that David Conrad's worries about 1.7 terabit attacks had much to do with RSSAC037 or this or anything else. That's an interesting data point, and I think we would have to think about it separately.

BRAD VERD: Yeah, it's certainly not what the metrics party was chartered for. I don't think, while that's a concern out there – it's been stated. We've been asked the question. Nobody in their right mind is going to provide a system that does a stress test to find the limit of anything. I don't see that happening. Certainly, if somebody

was chartering to do that, I would have real reservations and concerns around it.

JEFF OSBORN: My intro to RSSAC was a four-hour lunch where Steve Crocker would talk about nothing else. So it was in the back of my head. So if that's off – sorry. I had to [bring] …

BRAD VERD: That's not in the charter. We're not building a system that does any stress testing. We're building a system that is going to provide the health of the system and hopefully the metrics to be leading indicators of issues that can be addressed for existing roots. Obviously, one that rolls into 037 I would expect to used to validate the service of a potential root.

WES HARDAKER: I want to reiterate what I think Brad or somebody said yesterday, that, when we set out RSSAC workshops in general, one of the goals is to give a problem space. If the needs change in the process of looking at the problem space and we find that the charter needs to be changed, it is just fine to go back and update the charter.

Responding more directly to Jeff, I think the goals of this is to be able to measure this system and possibly individually roots to determine if they're able to contribute to the goals of this system. One of the goals of the system is to protect itself against various forms of attack. That's why we have a correctness measurement. Forget the DDoS attack. We have a correctness measurement to make sure that the right data is getting out there.

Right now, you're right that there is a prioritization by some thoughts in mind in particular directions in the same way that the web latency is the only thing they care about right now. We have this notion that we need to optimize for the current problem. I like to think of a more true engineering approach. What are the characteristics of the system we want, regardless of the current hard topic?

RUSS MUNDY: To get to the question of, in the document, removing the approaches that say, "Make use of a local recursive resolver, " the idea of having whatever the probe devices are do their own resolution. Then you should get a more accurate and more consistent set of results over time. As you do software updates to the probes, then they would tend to change together, and it would be independent of somebody deciding to put a new release of a resolver locally. So that was a lot of what led Duane

and I to this decision, that it would be better to go this way. We wanted to talk to the work party to see if there was agreement with that or not.

DUANE WESSELS: I would say, at least for now. If we get some experience in a future revision of this, we may come back this idea, I would say. But I think, if we can narrow how much we're biting off here, we'd be more likely to succeed in some reasonable amount of time.

Okay. I'm not sensing a lot more discussion about that, so should we move on?

RUSS MUNDY: Yeah, I think so. Since there was no objections – I think there was at least one who spoke up in support of the ideas –I think we should go—

DUANE WESSELS: We can take it to the mailing list and make it formal there, yeah.

Ozan, can we go to the PDF file from John? Ozan's going to put up a little presentation that was put together by John Kristoff. John is an ICANN research fellow. He's working with RSSAC and with SSAC. He was tasked with doing some data analysis related to the

root server metrics. Since John can't be here and it's very early for him, I'm going to walk through his slides, I guess.

All right. Let's go to the next one. John was asked specifically to look at the RIPE ATLAS data and used that data and applied it to the sort of things that are being proposed in the Metrics Work Party.

This text might be from the statement of work that John was asked to do. These two items represent where it's at currently. He's got some software that will do these parts, will download the data, and apply some of these algorithms.

Let's go to the next one. These other three are work that's still yet to be done. In particular, one of the things that I'm interested is the last one, which is to understand how sampling frequency might affect some of the results. But we're not quite there yet.

I think, looking at the data, he has it to the point now where he has analyzed and collected multiple days of data. I'm not sure it's up to 30, but I do think it's maybe up to 15 days or so of data.

Let's go to the next one. John's code is on GitHub. Anyone can go and check it out. I think he was asked to requested specifically to limit this to the RIPE ATLAS anchor probes, which is a subset of all the RIPE ATLAS probes. If you remember some of the data that I shared yesterday, that was from all of the probes.

What John has here is from the 500 or so anchor probes. The anchor probes, I think, first of all, are different hardware. They have a little bit more processing power. I think they're also located more in data centers rather than in people's basements. But I'm not positive about that.

As I reminder, all the RIPE ATLAS probes do very frequent measurements or queries to the root servers. So those are available going back, and he can download those. He's got some numbers here on how many data points there are today, per day. Looks like he's got about three weeks' worth of data analyzed so far.

I did install this software myself. I was able to do a little bit, but I got stuck at one point because part of it require Python Version 3, and the system where I was working was not happy to install Python 3. So I got a little bit stuck, but it does look like good stuff. Based on my own experience with using the RIPE ATLAS data, the parallelization is a key point. It's actually quite slow to fetch the results one by one.

All right, next please. Here's a big table of the 13 root servers across the top. These columns may be a little hard to read, but the major columns there are for IPv4, UDP, then IPv4 TCP, the v6 UDP and v6 TCP. Within each one of those rows, there are numbers for 25ht, 50th, 75th, and 90th percentile. These are all

latency measurements, I believe. I don't think John has done anything other than the latency measurements yet. There's no availability – I'm not sure that, for example, the RIPE ALTAS probes can give us the correctness measurements. So this is just about latency.

You can see the numbers here. Looking at some of the rows, you can see quite a wide variation in latencies, especially at the low end. Again, I'm assuming this is from the whole three weeks' worth of data, all aggregated together.

Any questions about this? I think it's pretty straightforward.

Yes?

LARS-JOHAN LIMAN: Having had to deal with the problem, measuring latency of TCP is interesting. So when you say "latency of TCP," you probably want to … if you want to work with details with these numbers, you'll also want to know exactly how that measurement was performed. It's much easier wit UDP because that's more straightforward. But when you need to do the handshake and turnaround times and what-have-you, it becomes more interesting. I just wanted to make that caveat before anyone else outside this room tries to use these numbers for any hard purposes.

DUANE WESSELS: That's a fair point. We can go and look at the RIPE ATLAS documentation and see if they provide that level of detail on how the TCP measurements are actually done. I'm not sure it's there. Maybe even Kaveh knows is the code is open-source. I don't know if we can go read the source code, but we can try to find out more about the details behind the TCP.

LARS-JOHAN LIMAN: I don't see that necessarily for here and now, and not even for this discussion. But if we want to drill down further into how this is exactly going to be performed in the future for these measurements, we may want to specify that.

Also, I would like to add that you asked before whether the RIPE ATLAS could be used for correctness measurements. I think it can. If you just seed it with a good query to send, you can probably use it for correctness as well.

DUANE WESSELS: Okay. Yeah, I wasn't sure if it gets the records/signatures all the time, but, yeah, we can check that.

To your point about TCP, Liman, I remember you said something similar in a previous meeting. The work party document does try

to be pretty specific about how to do a TCP-based measurement. So you should maybe take a look at that and see if it meets your expectations. For example, it specifically calls out that you start a timer when you send this in and you stop the timer when you get the response and you ignore the fin part of the TCP.

LARS-JOHAN LIMAN: Fair enough. I will read that. That said, it also depends on the speed of the probe computer.

KAVEH RANJBAR: Duane? For the correctness, I want to add here that it's possible to do the correctness. For the specific measurements, the TCPs [inaudible], but if you want to measure parts of it before the [scene] on all of that, we can also do that. We can change the code.

For the record, also the measurement code and how it's done – that's open source. [inaudible]. Actually, I think the DNS is mostly based on the busybox. But anyway, the code is there, and I can send the link to that. It's on GitHub.

DUANE WESSELS: I thought it must be open-source, yeah. Okay.

ABDULKARIM OLOYEDE: I just want to ask, for the percentile, of the probes that were considered on the 20th percentile for Root 1, are they the same for Root 2/Root 3?

DUANE WESSELS: Yes. I would think so. What I think John is doing here is he's taking 500 probes, all of the measurements that he gets from those probes, to Root 1, and then calculating these percentiles. So it's all of the same probes, except maybe there may have been a time where a probe didn't get a response – a timeout or something like that. But for the most part, it's always the same set of probes, yeah.

Let's go to the next slide. Here John has just a bunch of time series plots for each root server … this is hard to read. I'm not sure which one is which. They're not labeled, but we can assume one of these is v4 UDP and the other is v4 TCP and so on. Maybe we've got UDP on the left and TCP on the right, v4 on the top and v6 on the bottom. That would be my guess but they're not labeled.

This is … let's see. This looks like about maybe a couple weeks' worth of data. In these graphs, you can see that, clearly, there's a lot of clustering and overlap in the bottom, that the data points are a little bit thick so you can't really see any details down there. But some of them, you can see that there are the occasional outliers, but for the most part, it's very dense down at the bottom.

And you can see a lot of horizontal lines, indicating a lot of consistency, I think.

WES HARDAKER: The black box is almost measuring the 95% mark.

DUANE WESSELS: Probably.

WES HARDAKER: It's be better to do 25-, 50-, and 75-lines or something.

DUANE WESSELS: Right. Or to use maybe a little bit smaller dots or something like that to see more of the structure here. But we can take that feedback back to John and revise these graphs.

OZAN SAHIN: Paul Hoffman has his hand up in the Zoom room. Paul?

PAUL HOFFMAN: Thanks. If we could go back one slide, I wanted to point out something for a future discussion. Just looking at the first row, v4 UDP, 50 percentile – because, as Duane said at the beginning of the meeting, percentiles are probably more useful here.

As we get towards the discussion of thresholds, if we get towards the discussion of thresholds – I know that certainly some of the people want to get there – please do note, looking across 50%, and assuming this is milliseconds, a bunch of the roots are in the range that we often talk about liking, which would be, say, 50 milliseconds or less. And a bunch of them are well above that.

So, as we start thinking in terms of thresholds, we definitely need to think in terms of percentiles because, if you look at the 95% row in that first row, there's a lot more similarity, although some, like Root [16], still seem to be significantly faster than anybody.

But, at the 5oth percentile, which I think is a more understandable percentile, there is not as much diversity as you would think. That is, there's a lot of low numbers. And then a jump to a lot of higher numbers. I'm not saying we should use these table at this point, but, as we get to the thresholds discussion, this is actually very illustrative of how difficult it's going to be for us to say what is acceptable and what's not. Thanks.

RAO NAVEEDD BIN RAIS: If you go to the next slide, please, this graph I'm not sure is a collection of all root servers or just the RIPE NCC one.

DUANE WESSELS: This is one root server.

RAO NAVEED BIN RAIS: One? Okay. Can we conclude in this – I think we need some kind of research there – that this block box -- for example, if we just line it there – can represent 95, 96, 97 percentile of the overall latency measurement? Because that we can use as a conclusion out of this. That may help us having a threshold later on. So I'm not sure what this represents right now.

DUANE WESSELS: I don't know either. We could ask that those percentiles be overlaid here so that they show up more clearly. It's probably fair to assume that the large black box is about the 95% or maybe even a little more. But, yeah, we should ask for more …And I would put lines on this graph to show exactly where they are.

John has provided a set of these graphs for each root server, so let's go to the next one.

RAO NAVEED BIN RAIS: Can I have another one, please? Sorry.

DUANE WESSELS: Yes, absolutely.

RAO NAVEED BIN RAIS: I remember a discussion yesterday where you raised the point about standard deviation versus this percentile. So I wonder if we have a study on this particular graph that shows us the difference between the standard deviation and percentile and what we are achieving here. That would be interesting to see.

DUANE WESSELS: Yeah. We don't have that right now, but again, that's something that we could ask John to add, for sure. I would expect it to be similar because both what I showed and what John is showing are from the RIPE ATLAS probes – a subset, but both are from RIPE ATLAS.

RUSS MUNDY: Just as a reminder to folks that this work that was undertaken for us by the ICANN fellow was as a result of discussions in the work party for folks wanting to get some insight into numbers that we might actually be dealing with. What is something that exists today that could be at least somewhat representative of what we're talking about. So the RIPE ATLAS data, since it's publicly available data, was selected to do it, and that's what we asked the fellow to go off and do: this producing of some results based on actual data taken from the RIPE ATLAS anchors, right? It was the anchors? Yeah.

DUANE WESSELS: Let's go through these some more. I think they're all kind of the same but different. Sometimes you see different patterns, but they all have that characteristic where the bulk of the measurements are down at the bottom. You see different levels of outliers up at the top.

Again, if you are so inclined, this code is on GitHub. You can run it for yourself. You can make some of the enhancements that have been proposed here. You can add the standard deviation stuff. Then I'm sure John would love to get a pull request if that's the way you like to work. He would love some help on this. Otherwise, I know he's going to keep working on it and revising it. Hopefully, in the future, we can also have some data from some of the other metrics, not just latency.

KAVEH RANJBAR: Duane?

DUANE WESSELS: Yes?

KAVEH RANJBAR: One additional point which might be useful. Right now, this is not yet production, but we are internally testing. We have tested

uploading ATLAS [later] to the Google public cloud, both on BitQuery and BitCable to reduce time between sites. Especially with BitQuery, basically we can come up with some queries which do exactly that. That will be that and basically live. My assumption is that, by the RIPE meeting in October, that would be a production service, which might make this much easier because then you can basically do a simple bit query and you will get live results.

DUANE WESSELS: That would be cool. Is that a reason for me to go the RIPE meeting now?

RUSS MUNDY: Quick question, Kaveh. Do the RIPE probes or anchors yet do DNSSEC validation? Can you craft a query that will then do DNSSSEC validation?

KAVEH RANJBAR: The validation you can test, but crafting a query, like if it's non-standard, I don't know. There are some fields that we can play with, but there are some limitations as well. We already had many cases where people wanted to able to set that bit or that specific [r-type] things like that. That we can add. But I can see in your link

what's possible today. If there are other things, we might be able to add.

RUSS MUNDY: Thanks.

LARS-JOHAN LIMAN: Having worked a bit with the ATLAS system, you can set the DO bit to get the data, but I don't think you can get the probe to calculate the actual validation.

KAVEH RANJBAR: That's correct.

LARS-JOHAN LIMAN: So you will retrieve the data and you can do the validation yourself, but you cannot really force the probe because it doesn't have a whole lot of CPU.

KAVEH RANJBAR: That's true.

DUANE WESSELS: For the graphs that I was showing yesterday, those were based on the RIPE ATLAS measurements that were doing hosntname.BIND

queries. In that case, you don't get DNSSEC because that's not a validatable name. If I remember correctly, I think I chose those for a couple reasons. One because I think they were more frequent than maybe the SOA queries, and also I wanted the hostname.BIND data. But I'll have to ask John. These might be from the SOA queries, in which case you would probably get the signatures along with them.

LARS-JOHAN LIMAN: Yes, you will get the signatures if you set the DO bit. I believe that all the current roots now support the NSID option. You can set that request for that option as well. So you can get away from needing the hostname.BIND thing data and just rely on the SOA query or whatever you want to query for.

DUANE WESSELS: But what really matters is – these are standard measurements that are being done – how those measurements are being configured. They probably do set those fields, I'm guessing.

LARS-JOHAN LIMAN: Fair enough.

KAVEH RANJABR: I think they do. I can check and I will get back to you. The other thing is that all data is stored so we have the raw response. That is stored, so we can also look back and validate it if we want.

DUANE WESSELS: Another interesting thing I know that the RIPE ALTAS probes and anchors do is pings and traceroutes. We haven't done any analysis of that data at his point, but we could consider that as well. It might be interesting to look at that. But for the most part we just asked John to look at the DNS measurements and how they can be applied to the Metrics Work Party.

I think we're wrapped up on that. How much time do we have left?

RUSS MUNDY: We have 40 minutes.

DUANE WESSELS: 40 minutes? Wow. Okay. Who wants to talk more about probe locations?

Nobody. Didn't think so. Ozan, can you maybe go back to the Google Doc then? I've been leading a lot of the discussion, Russ. Do you want to lead any parts? Otherwise, I'll go through some of the comments that we have currently in the document.

RUSS MUNDY: Yeah, let's just go over some of the comments that are currently in the document.

DUANE WESSELS: Okay. I'm not in the Zoom room, but somebody else – I think Carlos – can …

In the room here we're just waiting for the screen, the display, the change to the Zoom room.

OZAN SAHIN: The document is showing in Zoom room, but we are trying to get that in the—

DUANE WESSELS: But not in this room.

OZAN SAHIN: Yeah.

RUSS MUNDY: Okay – oh, yes. I do have the Zoom room open and, yes, it is in the Zoom room. Thank you, Ozan.

Well, happy to see there's a number of additional comments today and the last day or two. So this is good.

Wes, question for you. I see some comments from John Heidemann. Do you know if he's planning on joining the call at all?

WES HARDAKER: It is 3 A.M., so no. Somebody asked me to bug him. I don't remember who the somebody was, but they were hoping he could review it. So I did that yesterday, and he submitted a bunch of comments. As well, I submitted a bunch yesterday. So there's probably a lot more since the last time you looked.

RUSS MUNDY: Okay. So we will put notes in as appropriate comments for folks that aren't able to be in this particular meeting.

WES HARDAKER: That's fair. I'm able to channel him fairly well a lot of the time, so, depending on the comment, I might be able to figure out what he's saying, if you have questions.

DUANE WESSELS: All right. We've got it up on the projector here in the room. I think what I'll do is walk through the document and point out certain things. We can have discussions about questions that folks have raised or not. We could leave them for later discussion.

Can you scroll down to Page 5, the introduction? I think that's where the first comment is.

This is something that Fred had just mentioned a while ago. I think this is a good catch. The very first bullet point here talks about how this metric is relating to SLAs. It doesn't say SLAs, but essentially it's talking about SLAs and performance. Maybe we also want to make a note about the fact that these measurements can be used for other purposes – research and the other thing you talked about, which is –

UNIDENTIFIED MALE: [inaudible]

DUANE WESSELS: Yeah. So I think that's a good catch. Scrolling down a bit to the background and scope, there's a definition of instance. John has a comment that this may be better as an Anycast instance. I don't have a super strong opinion about that, but I would note that this entire thing is really directly quoted from the RSSAC lexicon. So this is the definition of instance from the lexicon. There is also a definition of Anycast instance in the lexicon. So maybe we need both. I don't know.

WES HARDAKER: I think adding …

[KARL REUSS]: As [we] could say, the lexicon definition of instance is a little ambiguous because in the same – since it refers to it as a server and as a location. So is it an IP address server or is it a location?

DUANE WESSELS: So Karl is offering to chair the next work party to update the lexicon, it sounds like. No.

WES HARDAKER: John I don't think is aware of the lexicon document because the next comment actually down from him was similar. He was suggesting a different name. I responded to that comment that this is actually a direct copy. It's not actually from the lexicon.

DUANE WESSELS: Yeah. We tried to do that, although I do note that there is one – is the root server definition? One of them we specifically modified or a little bit amended the definition from the lexicon. Which one was that? Ugh.

WES HARDAKER: You probably don't need to find it now, but I would make the suggestion that, if you did modify it, you somehow indicate that in the text.

DUANE WESSELS: Well, that's what I'm looking for. I thought I did that.

WES HARDAKER: Or, in the background of this, I would say that most of these were taken from the lexicon and that's the authoritative source for …

DUANE WESSELS: Is it root server? Yeah. So it's the definition of the root server. The second sentence of the root server definition says, "Use in this document differs slightly from the one in RSSAC026.

Then I see there's another comment here from John, which looks reasonable to me at first pass. So we can take that to the work party and see if there's agreement to accept that.

PAUL HOFFMAN: Duane, I'm going to jump in without raining my hand. Sorry.

DUANE WESSELS: Hi, Paul. Please.

PAUL HOFFMAN: You sort of jokingly pointed and said, "Oh, maybe we need to update the lexicon." If that's actually a desire, since this document is really the first in-depth use of the lexicon, I'd be willing to do that, just since I seem to be the lexicon dude on DNS for DNS [op], even if it's a small update. If we as a work party are finding things in the lexicon that are either unclear or really need to be update, we might be able to do a quick update on the lexicon, and I'd be willing to do that.

BRAD VERD: Hey, Paul. I know that updating the lexicon is in our queue of work to do. When we sat down – I forgot where; maybe it was Barcelona – there was a list of documents that we thought we should, one of which was the lexicon. So this is in the queue, but if you want to take the bull by the horns, I think that'd be welcomed.

WES HARDAKER: Paul, can you speak to, since you're, I think, in both the overlap between the IETF DNS lexicon and the RSSAC one – do you know how well they aligned? I figure the one that would know.

PAUL HOFFMAN: They are barely aligned. I don't think that the IETF lexicon actually quotes from the RSSAC lexicon. I could be wrong, but I know that it refers to it as, "Oh, look. There's another lexicon over there."

WES HARDAKER: All right. Thank you.

RUSS MUNDY: Do we want to take a general action to go through the review of the terminology separately for all of these and work at it, especially with looking at the possibility of revision to the lexicon as needed? Because right now in our wording in here we tried to identify in the individual ones where there is a difference. And then make a determination if working it to where they wouldn't be a difference so we wouldn't have to have that amount of text in this.

DUANE WESSELS: Yeah, it sounds reasonable to me. I don't think we want to make one necessarily dependent on the other but, yeah, we should keep them updated and in sync.

All right. Let's scroll down a little bit further, Ozan. I see a lot more comments from John about … I'm going to suggest that we skip through some of these terminology definition things for now and

move on to maybe some more meaty topics. I think John is reading this with fresh eyes and finding a lot of things that we have missed while we were writing this, and these can be easily addressed.

Can you go to Section 3? Here, for example, Liman, is where this stuff about TCP is located. It's actually at the start of the next page. I see there's some comments about this section. Again, I think this is just minor wordsmithing. One thing to note, though, is that there is a part here about TCP Fast Open. What does John have to say there? "The rule here seems odd. There are many details and options to TCP. Shouldn't this be left up to the operator?"

The reason that we mention TCP Fast Open at this point in the document is to say that, in order to be consistent, a probe doing a measurement over TCP should not use TCP Fast Open to be consistent over time.

I think John's comment is more about the service side of TCP Fast Open. Of course, in order for that to work, you need support on both the client and the server sides. But what this document clearly states is that, if you are doing a TCP-based measurement, you should not use TCP Fast Open. If you did use TCP Fast Open, then essentially you would be measuring the same latency as UDP, I think. That's what it's designed to do.

If anyone has opinions on that, I would welcome them at this time.

Scrolling down a bit, Section 3.3 is really something that we've carried almost verbatim from a very early work party document. This really gets to this sort of thing we spent a lot of time talking about today. So I expect this section to change a lot in the future. This may where we introduce some notion of near probes and far probes if we settle on that.

The next one talks about use of a probe's local recursive name server. Again, based on the discussion we had just a few minutes ago, this might be going away, I would say. We may eliminate that method of measurement from the document.

Section 3.4 talks about spoofing protections. This was added, again, to account for the case where – in my mind, we had probes that were far away. There may be stuff in the middle that's interfering and a probe should do everything in its power that it can to try and ensure that there's no spoofing going on.

If we're going to instead consider doing some of these correctness measurements from nearby probes, then this becomes maybe less important.

WES HARDAKER: Can we talk about the other aspect of this, which is whether or not we should be using a recursive resolver? What we're pointing out talks about – I made a comment about this as well. In my mind, we have a number of options, one being we use whatever recursive resolver was handed to us by [DayHEP], which the document currently throws away, which is fine, and the second one being using a resolver that we know is running on the local host. But that's hard to make sure that's always happening. If we control everything, then I guess that's the case.

If we're deploying, say, a full OS or at least a Docker container that contains both something, the other option is we actually build in a recursive – we put it in the specifications – library, and there's a number to choose from directly into the probe software itself.

Any time you are doing measurements, we need to know exactly how it is being measured. Knowing what recursive software is in use – be it a separate, the same process, or a different host – has to be known in order to actually determine results from a measurement.

DUANE WESSELS: Are you talking about measurements via a recursive nameserver?

WES HARDAKER: Yeah because …

DUANE WESSELS: But Russ and I just proposed eliminating those.

WES HARDAKER: Okay.

UNIDENTIFIED MALE: Hear, hear.

DUANE WESSELS: So I think that's maybe now overcome by events.

WES HARDAKER: Okay.

DUANE WESSELS: But I was wondering if you were talking about the need for a recursive sever for purposes other than measurements, like to resolve names to report back to the central system or something like that. I think, in the interest of simplicity and further reasons, we would like to eliminate measurements via recursive servers at this time.

LARS-JOHAN LIMAN: Strongly supported by Liman.

RUSS MUNDY: There is one comment in there from John about the naming. We've been calling them probes. John suggests another term be used. I don't know what folks think about this, but it seems like this would be a good time to discuss whether or not we want to try to come up with a different naming besides probes. He says "vantage point."

Any support or further on thoughts on that? Wes, do you want to comment on why John might [inaudible]?

WES HARDAKER: I can channel him some. Vantage point is the typical term used, especially in academic publications, for where you're doing measurements from. It's typically used when measuring Anycast networks or things like that. So my guess is that that's where his thinking comes from.

LARS-JOHAN LIMAN: I was just going to ask for a motivation. I've just received it, so thank you.

RUSS MUNDY: Does anybody have any strong feelings one way or the other about changing the names? One of John's point that he, I think, is

making quite effectively is that probes is a very widely and heavily used, heavily loaded, term. But what other terms might be good? Vantage point was what he suggested.

Anybody want to comment one way or the other one that?

DUANE WESSELS: Initially, I do like vantage point. I think that has some advantages, but I would need to think about it some more because I think sometimes we also talk about a probe as a device, like you want to hire someone to build a probe. They're not building a vantage point. They're building a probe. So I like it but I think we need to think about it a little more, maybe.

RUSS MUNDY: After seeing his comment and only thinking about a few minutes, what strikes me is that perhaps the term "test point" might be better for what we're doing. But this is – Fred, did you have a thought?

FRED BAKER: Well, I was just coming back to my thing about, gee, what if somebody wants to test their scenario and wants to use these metrics to do so? I could imagine a piece of open-source software

that people could download and do these tests. Is that a probe? Is that probe software? Whatever that word is.

LARS-JOHAN LIMAN: I'm with Duane. These are slightly different things. The probe is the unit, and the vantage point is the logical place from which you are making the observation. So it could be that we actually need to use the different terms in different places in the document because we actually mean the different things.

I don't think we should introduce a third term like test point because that's not anything I've heard anyone proposed being in use elsewhere. Then it won't benefit us to create a third one. So I would encourage Duane to look at the document and see which ones fits because I'm quite [inaudible] with the distinction.

DUANE WESSELS: Thanks. Liman, I want to actually get back to you. Something I think you suggested in our April workshop was to use 6-0. You said "half-jokingly," I think was the phrase. So there's a little to-do item here in the document about that.

I question whether it's realistic, but I guess I wanted to give you a chance to argue for it some more.

WES HARDAKER: Do you want another argument before you argue back? The comment that I put in, since I doubt anybody's read it, is that the problem is that the instance you use something like6-0 you are throwing an error into the measurement because you are measuring the service in a way that the Internet users who you are doing this on behalf of won't be doing. So you're affecting measurements, like latency and other stuff, significantly, by adding cryptographic overhead.

For correctness, it probably won't mater, but if you're doing correctness, you already have a cryptographic measurement that you're making use of. So I was pushing back against the potential.

LARS-JOHAN LIMAN: Thank you, and fair enough. Yes, I did say it half-jokingly. I would argue though that 6-0 and DNSSEC provided different types of authentication.

WES HARDAKER: They do. True.

LARS-JOHAN LIMAN: So if you really want to authenticate the server as opposed to the data, you need something else than DSSEC. Again, I'm not sure that's necessarily, but there is a slight distinction.

Toss it away if – I'm quite with you both, Wes, saying that that's not going to be used by real users, and with Duane saying that it's probably not realistic because it's not implemented. So absolutely. Fair enough. Throw it away.

DUANE WESSELS: As to my understanding, it would require putting key records in rootservers.net zone, for example, which opens up a whole complication.

LARS-JOHAN-LIMAN: That might, yes, well be the case. So I fully agree. Let's drop this.

DUANE WESSELS: All right. Next, Ken, you suggested this nice block of text about measurements based on user experience. I think this maybe would fit well with what Fred was saying up earlier in the document.

KEN RENARD: Yes. A lot of this was captured by the discussion this morning. If this text is usable anywhere, great. Glad to contribute. But the goals, the bulleted list down there, was more geared towards how do we validate a measurement? How do we detect spoofing and debug, which is a slippery slope maybe outside this document?

If we're going to capture the discussion this morning of the near/far, this could be completely replaced by that.

LARS-JOHAN LIMAN: But we really want to measure the resolver experience rather than the user experience, right?

DUANE WESSELS: Yeah. I think you're right. When I thinking "user experience," I'm probably really thinking "resolver experience," as them being close together. But they're not always necessarily close together.

All right. Let's scroll down a little bit further. There is a section here that describes random non-existent query names. I think that, again, if we're going to eliminate measurements from – well, actually that's not true, now that I think about it. There are some measurements for which we wanted a name that was guaranteed tor result in an NX domain or to not be in cache or something like that. Some of those measurements were done via recursive nameservers, but now that I remember, not all of them are. I think the correctness measurement now uses this random NX domain algorithm to generate a name that would not be a delegated name.

John has a question. "Why not add a fixed identifying string like test?" I don't have a strong opinion about that. To me, as it is, it

serves the purpose. I don't know what adding "test" does, other than make it identifiable.

WES HARDAKER: From a data analysis point of view, being able to filter stuff that you knew what it was, whereas this identifier is not labeled, essentially. Labeled data is always a hard thing to come by, and, if any time you can hand me a label so I can clean my data when I'm trying to do analysis of it, that's highly helpful. So, for people analyzing DITL, having something like "test" in the name or maybe something specific to this service so we can actually look up every – all of these queries came from the root monitoring project. So we can either look at them specifically or throw them out specifically.

UNIDENTIFIED MALE: I am in the queue.

LARS-JOHAN LIMAN: Okay. Two comments. One is, if the string is known, someone will game the system. That's a given. The second one is that – and this is jokingly – you could ask Chrome for the data because they send us three of these queries every time they fire up.

WES HARDAKER: Bite your tongue.

DANIEL MIGAULT: One thing is, instead of—

UNIDENTIFIED MALE: We can't hear you [inaudible]

DANIEL MIGAULT: Okay. So, instead of using a random name, why don't we have just a name that is not in the root zone?

UNIDENTIFIED MALE: [inaudible]

DANIEL MIGAULT: Yeah. [I think so on here].

LARS-JOHAN LIMAN: Because it will get stuck in the cache, and next time you ask, it will come from a cache.

DANIEL MIGAULT: But there is no cache from the prob to the RSO.

LARS-JOHAN LIMAN: Yes there is. In 25% of the cases – I'm just pulling a number out of the air – there's something in between, yes: firewall—

WES HARDAKER: If we don't use a recursive resolver and we control the probe, then that's not true. One of the other issues that I had somewhere else in the document is – the point that I brought it up was I think we say that there's a four-second timeout. If there is a man-in-the-middle box that's doing what you're talking about, a forced DNS proxy, it could return a timeout faster than that. So there's other cases where what are we going to do about middleboxes being in the way? Or do we want to, somewhere higher in the document, say probes shall never placed in any location where we believe that there's a middlebox or that, if we can determine that, we [will] stop using that probe or something?

LARS-JOHAN LIMAN: I thought that this section was actually partly to identify the situations where there is a middlebox in between. I will argue that it's unavoidable that there will be middleboxes in the way. We absolutely must design a system that can handle that, or at least tell us when it happens. So hoping that there is nothing in between is a hope in vain. We just have to live with that. We need to design this system to be able to tell us if there's something in

between and, if so, how it behaves so we can take that into account when we analyze the results.

DUANE WESSELS: Daniel, the reason to not use a fixed name here is that, for the correctness measurement, the current proposal is – so it validates DNSSEC signatures. With certain probability, you would query for an existing name – com, net; whatever. For some other probability, you would query for a name that does not exist so you can validate the NSEC 3 and the NSEC 3 signatures. You want it to be different so that you hit a different area of the root zone at different times. If it's fixed name, you're always going to get the same response back for the correctness metric. So that's one reason that it's not a fixed name here. Does that make sense?

BRAD VERD: I would question why we're engineering to identify middleboxes and the engineering to account for changes that they might inflict on the response.

LARS-JOHAN LIMAN: Because that may affect both the delay and the correctness of the data.

BRAD VERD: Right, but wouldn't it affect the delay and correctness on every server coming from that probe? So it wouldn't show a bias towards any server. It would affect every data point.

LARS-JOHAN LIMAN: I'm not going to take that for granted because that operator really hates Netnod. So they inflict pain on the queries going to us. There's so much crap out there. And I just had contact with another guy who ripped Netnod out of the hints file. He had reason to do so, but people do the strangest things and we need to understand what's going on when we do the measurements.

[JEFF OSBORN]: There's ATLAS data that shows that individual RSOs are treated differently.

DUANE WESSELS: Section 3.6 is a little blurb about Anycast. This, again, may need to change if we settle on using near probes or at least probes that are directly associated with an individual root server instance because in that case we are in fact measuring all the instances or some subset of instances, whereas what this paragraph says is that measurements are only being directed towards root server addresses and not really caring which instance is serving the response. So just to point out that this may change.

DANIEL MIGAULT: But I don't think it really changed because we're not really measuring how effective BGP is.

DUANE WESSELS: No, we're not. That's right. We're not measuring BGP. That's not one of the goals here. If we're considering a system of vantage points next to root servers, then the idea is that you are specifically targeting that root server instance close to the vantage point. But again, yeah, not trying to see what BGP is doing. Or probably not.

RUSS MUNDY: I have a question for Fred about the comment you have here in the Anycast section. Are you suggestion then that we should have something in the document that at least would somehow permit identification of what instance is being hit? One of the things we've tried to do so far is to, from the metrics perspective, just say we're not treating Anycast instances at all and [our] RSO operation is the level of granularity we're going to. So are you suggesting we change that in some manner here, Fred?

BRAD VERD: Think of a case.

FRED BAKER: I'm thinking of a case. In the cases that, here I am, doing some measurement and doing whatever it is that I'm doing, it becomes useful to know which instance I'm looking at and to know whether I've hit all of the instances if I was trying to do that or this kind of thing. I am thinking larger than, am I running an SLA here? I'm thinking in the research case. It just seems likely that somebody will sooner or later want to know, here I am, looking at something: which instance is it? Now, I don't know that I'm changing the metric in any way, but that seemed like a question that's going to come up.

BRAD VERD: Yeah, I can guarantee you that'll be a question that comes up because, in the event of an SLA or a breach and you go to the root server to say, "You're failing this test," they're going to say, "What are you hitting?" so I can go try to address it.

DUANE WESSELS: Yeah. We previously talked about this in the context of the correctness metrics, that, if there was any situation in which it was not 100%, then the report should also include as much details about the exact probes which probes which caused it to fail, which [NX] measurements that caused it to fail.

BRAD VERD: Like if the probes could do a traceroute or something there.

DUANE WESSELS: Something like that, yeah.

FRED BAKER: Yeah. I absolutely agree that, in many cases, we're simply trying to describe the label, but it seems we need to think a little bit outside that box.

RUSS MUNDY: Well, as someone who's basically a researcher, I'm very much in sympathy with that. But as one of the Co-Chairs of the group, I want to have us get completion of our document here.

FRED BAKER: Well, you just heard from an operator.

RUSS MUNDY: So I'm not sure how many and what type of changes or additions or modifications it would take, but I think it's something we need to think about. Certainly, if there are identifiable ways that already exist and are already defined that can be incorporated

here, that's the thing to do. I'd hate to have the work group to try create something on their own to identify an instance.

DUANE WESSELS: Yeah, I don't think we need to do that. I think there are ways to do that that are straightforward.

I know we're – I think it's almost quitting time, right? We're right up against that … I do want to also point out one more thing for people that have the document up and are willing to keep looking through it. One of the things that's changed since the last we've presented this is we've adopted a little bit of the style from the IANA reports, so, in each of these sections, take a look at the example report or the example metrics and see how they're listed there and see if that meets with your approval or not. Give us feedback on that.

Otherwise, of course, keep scrolling the document and look through comments that people already made and reply to those. We'll take all these up next time the work party meets and try to make progress on this.

RUSS MUNDY: From what Fred said at the beginning of the meeting, you expect the modern resolver behavior work party to be less than the full

allocated time? And then you're going to push that back to metrics? Was that the plan?

FRED BAKER: Well, yeah, that's my plan. Let me comment on this afternoon, which I was about to do anyway. At 13:30, we're coming back and we're talking about resolvers. I have a hunch. Paul Hoffman is quite free to correct me, but I bet we're not going to use 90 minutes for that meaning. By all means, let's continue with metrics in whatever time remains available.

Then, at 15:15, we have the RSSAC meeting, which is going to be basically our July meeting. That's going to happen this afternoon.

BRAD VERD: And that's not going to use 90 minutes either.

FRED BAKER: Yeah. So if we have time available at the end of that, then [we go].

PAUL HOFFMAN: Yeah. As I said in a message to Fred offline, it's very unlikely that we would use anywhere close to the 90 minutes. So, for those folks who are on the call who care about metrics but don't care that much about modern resolver behavior, I would suggest

sitting through the modern resolver behavior stuff so we can hop back into metrics.

RUSS MUNDY: Lunchtime.

UNIDENITIFIED MALE: Lunch.

[END OF TRANSCRIPTION]