
MARRAKECH – RSSAC Work Session: RSS Metrics (part 1 of 2)

Wednesday, June 26, 2019 – 09:00 to 10:15 WET

ICANN65 | Marrakech, Morocco

FRED BAKER:

We'll give it until five after, and then we'll blaze on.

Okay. All fine. Let's talk about the agenda for the day. Just a minute.

So this morning, we have dedicated to the metric discussion. We have a slot this afternoon which is, at least on the agenda, dedicated to the resolver discussion.

I'm pretty sure that resolver discussion will not use the entire time. So what I expect to do is continue – assuming we need it – with the metric discussion once the resolver discussion is completed. And then in the last work session of the day, I believe we have the monthly RSSAC meeting, which will actually, in terms of calendar, be our July meeting. We will skip the July meeting in view of the fact of having one in here.

So let me point out that they're all open meetings. The first three meetings, the metrics meeting and the resolver meeting are caucus meetings that happen to be occurring in the RSSAC place here. So you people that are caucus members and are sitting out

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

in the boonies, please feel free to come up to the table. You're as much part of this as we are.

So with that, I'll turn this over to the metrics chairs who are sitting somewhat to my right.

CARLOS REYES: Could we do an attendance before we get started? Thank you. So maybe we'll just start at one of the ends here. Thank you.

KENNETH RENARD: Ken Renard, ARL.

KARL REUSS: Karl Reuss, University of Maryland.

DANIEL MIGAULT: Daniel Migault, liaison to the Internet Architecture Board.

DUANE WESSELS: Duane Wessels, Verisign, one of the metrics work party co-chairs.

RUSS MUNDY: Russ Mundy, SSAC liaison to the RSSAC and one of the metrics co-chairs.

BRAD VERD: Brad Verd, co-chair, RSSAC.

FRED BAKER: Fred Baker, ISC, co-chair, RSSAC.

CARLOS REYES: Carlos Reyes, RSSAC support staff.

OZAN SAHIN: Ozan Sahin, ICANN support staff, and on the remote, we have [inaudible] in the Zoom room, Kazunori Fujiwara, Keith Bluestein, Paul Hoffman, and Terry Manderson. And Brad is already in the room.

HIRO HOTTA: Hiro Hotta from WIDE plus JPRS.

LARS LIMAN: Lars Liman from Netnod.

UNIDENTIFIED MALE: [inaudible] RSSAC caucus member.

RYAN STEPHENSON: Ryan Stephenson, Defense Information Systems Agency.

WES HARDAKER: Wes Hardaker, USC ISI.

SHINTA SATO: Shinta Sato, JPRS, RSSAC caucus.

UNIDENTIFIED MALE: This is [inaudible] from Computer Network Information Center of CAS.

UNIDENTIFIED MALE: I'm from China Academy of Information Communication Technology.

DUANE WESSELS: Is that everyone? Okay, so we're good to go. So again, I'm Duane, I'll be probably leading the discussion for this first session at least. Before we get started, I wanted to say a little bit – since we have the caucus members, we have more participants here, I want to say a little about what we talked about yesterday.

We had some free time in the agenda yesterday, so we did a little bit of metrics discussion then. We kind of talked about two things. One presented some graphs of actual data showing why

percentiles make a better representation of precision than standard deviation for the types of distributions that we see in existing root server traffic, and apologize that this wasn't sent out before, but I can send these slides out to the caucus mailing list so that everyone has them.

That discussion went pretty well, and I think everyone agreed that we should use percentiles and avoid standard deviation. Then we also talked about the BPQ section of the work party document. Specifically, what we previously agreed is that the work party should come up with a method for determining or estimating BPQ for the RSSAC 002 data.

So there's a little bit of text written about how that might work, and I just wanted to make sure people were aware of that. And again, I think that's settled now.

By the way, I'm not in the Zoom room, so I guess Carlos or Ozan, if you see somebody that has a question, please interrupt and let me know. I won't be watching.

Okay, so today, we want to tackle some bigger topics. And to start off, I'll just refresh everyone's memory that in the work party document, most of the things we focused on, there are sort of four metrics: availability, latency, freshness and correctness, and those apply both to the individual operations and the whole root server system, although we did previously agree that we should

strike freshness for the whole root server system. So in that case, we just have the three.

And previously, we'd also sort of talked – there had been a suggestion that when we design our measurements and our metrics, they should maybe be designed in a way that they could utilize an existing measurement platform, of which RIPE Atlas is probably one of the best examples.

But sticking to that idea, I think we encountered some problems and some difficulties, namely that when we have probes that are very far away from the servers that they're measuring, there's a lot of network stuff in-between, a lot of chances for things to maybe go wrong or be meddled with.

Previously, I sensed some hesitation from some of the root operators that they would be asked to be held to service-level expectations for parts of infrastructure that they can't really control.

I think one of the areas that this is really a problem would be like in the correctness metrics, because there was agreement that we should strive for 100% correctness all the time, and that means that some actor in the middle, either intentionally or unintentionally, is able to interfere with these measurements or DNS traffic in general could impact those type of metrics.

Then I got to thinking about, well, if we were designing this from scratch, if we had no limitations and no previous expectations, how would we do it? And I think for some of these metrics, we might use measurement points that are much closer to the servers. And the correctness measurement would be one of those, I think.

So on the last metrics call, I'd put this idea out there that maybe we have two types or two sets of probes. We have probes that are near to the servers and probes that are far from the servers. And you might have different measurements being done from both sets of probes.

So that's really the topic of discussion today, and I think this is also sort of all tied up with some other things that we need to talk about, which is the idea of metrics that are specific to SLAs or designed specifically for SLAs and metrics or measurements that are not really based around SLAs that are more just sort of informational.

Fred, you had your hand up a second ago. Do you want to jump in?

FRED BAKER:

Well, yeah. You mentioned that there would be different metrics in the different places or different measurements being taken in

different places. And I suppose one is the correctness, is the data correct?

And here I'm replaying a conversation that you and I had the other day. It seems like what we're looking at is an SLA discussion of the RSO, is it doing its job? And separately, research question of what is the user experience, and we're using actually, I think, the same metrics to do both, but they're in different places with different interpretations. Is that what you were getting at in your comments a moment ago?

DUANE WESSELS:

Yes, I guess, essentially. There have been no proposals at this point to have like something called a practice measurement that is done on the near probes, and some kind of different correctness measurement that is done on far probes.

I think the measurements and the metrics are simple enough that you can apply the same definition to both. It's send a query, get a response and validate it. It's basically that simple.

Now, if we needed to make them different, I suppose we could, but it would be better if they're the same.

So, again, to provide a little more thought into this, I think one of the things that we want to get out of the discussion today is if – well, first of all, if this is a reasonable idea, to have these two sets

of probes, and then if it is, let's talk about what types of metrics we would want to see from those two sets.

And for example, latency in my mind makes the most sense to do from far away probes. If latency from probes that are near to the server, while you can certainly do those, I think they become less useful and less meaningful, because in my mind, when I say near, I'm thinking like almost adjacent to the server.

So everyone would have almost the same latency. There's no differentiation there, and as a metric, as an SLA or an SLE metric, there's no real teeth there. Everyone can meet this type of metric. So that may be an example that we only want to do from the far away probes.

The other thing I want to sort of throw out there right now is, I think if you can imagine this nearness or farness as a slider that you can control, the nearer you get, it has certain properties and benefits, and farther away you get, it also has certain properties and benefits and disadvantages, and maybe there's some middle ground, maybe we need to adjust this so that we're not talking about far away in the sense of like the RIPE Atlas probes, which essentially at the user sites or in people's homes and businesses, but maybe somewhere in the middle, the network in the middle would be a better place for a far probe or a medium far probe.

Brad?

BRAD VERD:

Yeah, I'm sorry, I have to leave. I have a conflict. I'll leave for 9:30 and hopefully be back before this is over. But I did want to add some commentary here. I like the idea of near and far probes, but I've thought about them a little differently, and I wanted to drop this grenade and then leave. Which is not planned, it's just how the conversation worked.

But I kind of think about them – and I mentioned yesterday, I feel like maybe we don't want to build a massive monitoring platform. I think it should be reasonable, but small as possible, but still get valuable data. I think I said like top ten peering points is what I said yesterday. That would be the far probe.

And I see those probes, the far probes, as monitoring two different things. There's the health of the root server system, and then the health of the root server operator service, so essentially each letter operated by a root server operator.

The near probes, as I thought about it, this was going to get complicated as we've got to deploy probes by all the instances. And just as I started to think through it, my head started to actually hurt, and so I kind of applied what's in practice today for us, which is we from the TLD perspective or the gTLD perspective, we do self-reporting, so we have SLAs, we do self-reporting. Those are auditable. We save the data, we report on them. And it seems

to me like maybe the local probes, I kind of thought of them as maybe SLA metrics.

Now, I know this metrics discussion is not tied to 37, and I'm not trying to connect them here, but they are kind of connected in the sense that that is the trajectory we're on. So maybe out of the metrics work party comes, if we reach SLAs via the implementation of 37, that we could have a recommendation there.

And the recommendation could be a set of metrics that are done, that are self-reported as soon as – if you're collecting funding, then there are metrics to be reported, and it's self-reported.

So that was kind of my train of thought on how to do that without putting a requirement out there to build this massive monitoring platform. So that was one approach that I wanted to talk about. Unfortunately, I have to step out of the room.

Anyways, I'll leave that for the conversation. Happy to talk about it when I come back though.

FRED BAKER:

Well, you said a minute ago that this isn't necessarily immediately tied to SLAs and SLA performance. Actually, the opening sentence of the report says something about SLAs measuring performance. So it's not that far away.

But my real reason I raise my hand, I'm thinking about the definition of near and the definition of far, and Ken or Ryan, you might actually be in the best position to answer the near question. If I'm putting a probe in front of your services, if you're putting a probe in front of your services and reporting how well you're doing and so on and so forth, I believe you have a mandated firewall. And that has, in the past, impacted some of your performance.

Would near be in front of the firewall, or would near be at the server behind the firewall?

BRAD VERD:

Just really quick commentary, the reason I thought this through as far as self-reporting and comparing it to what happens at least on the TLD level for us is there was a bunch of comments in the work party about how can I be responsible for a packet going across the Internet and the latency and all these different things? Which is reasonable.

So if you look at the gTLDs, I believe the SLAs are basically when the packet enters your network. So for us, we have to monitor as soon as it hits our border router, through our network, out the border router. Those are the metrics that we're responsible for, because we're in control of it.

So in your case, Ryan, I would assume you're in control of it, you would have to accommodate for that firewall. That's how I think about it. But I'd love to hear other thoughts. That's why I was trying to think of the self-reporting aspect, because of all the commentary about how can you have a metric that goes across the Internet that I'm responsible for?

RYAN STEPHENSON

So with G root, it's buried down into a separate ISP that we connect to another ISP for Internet service, and to the global Internet. So with that being said, putting in a probe or just any type of other device that's not approved through cybersecurity or information [assurances,] basically just hooking up any type of equipment, considering the fact that we can't even plug our personal cell phones into or laptop for power, that's – the definition of near is going to be somewhat relative, I think, for each root operator.

For us, the best definition of near would be either in a geographic area where our servers are located, or possibly at the ISP peering point, that is if we can coordinate with our ISP to say, “Hey, can you just put a probe within this particular peering point where we connect to?”

So for that, without giving away too much of our internal architecture, that would probably be ... so for us, that's where

near is going to be, really at the ISP outside the firewall. And yeah, there is a firewall, actually. You could actually google it at the ISP. From the ISP then, we go directly into G-root.

So for us, near is really going to be at the ISP level, and that is if we can coordinate with the ISP. We may have to – for example if the ISP says, no, we don't want to hook up a RIPE probe or whatever probe is going to be used, then we would just have to look for something in that local geographic area.

So, our definition of near is going to be different than, say, root X or root Y, or Netnod. So that's a very good question about the definition of near, and we do have to kind of keep that into term that it's going to be subjective for each root operator.

FRED BAKER: Please go on. But it sounds like then each operator is going to have to say what they mean by near.

KENNETH RENARD: So following on what Ryan was saying, in my mind, my concept of near and far was much more extreme. The top ten exchange points, that's sort of near to me. My concern was more far away, 15 hops behind an ISP over ISDN lines and 2600-BOD modems.

But going back to, Duane, what you were saying, what you opened with, that very accurately represented what my concerns were. So thank you for that.

Kind of combining these discussions, mostly for correctness where you have potentially lots of middle boxes, yeah, if we had the entirety of atlas infrastructure, atlas measurements, if you just threw away “far” probes for the purposes of correctness, for the purpose of SLA enforcement, that seems very reasonable to me.

BRAD VERD:

Can I ask a clarifying question? You said a top ten peering site would be near to you. Would you be comfortable being held to a correctness SLA from measurements made at those peering points?

KENNETH RENARD:

I'm open to pretty much anything here, but when I have these concerns, those top ten peering points were “near.” And yeah, I'm less concerned about that. I'd still consider that rather near.

LARS-JOHAN LIMAN:

Two comments. One, you made the comment I wanted to make, is that I think the important thing, for at least the near part is that

we document for each operator how and from where the measurements are made so that there can be a judgment. If the numbers turn out strange, there can be an explanation for that.

My second comment is that I worry about the outliers. If you have a large system of probes and measure from many different points, I want to cut off the tails in the beginning and the end so that we don't have a single point making a huge impact on the measurements and the results.

BRAD VERD: Yes, but the correctness one is sort of the extreme case because we said 100% correctness. Are you talking about latency?

LARS-JOHAN LIMAN: I'm talking about latency, yes. Correctness is a different issue, and I'm with you entirely. And you could measure correctness from just about anywhere, but latency – and even actually possibly reachability. But reachability is more complicated than latency.

BRAD VERD: Go ahead, Karl.

KARL REUSS: It seems like correctness for the near case might be a little problematic in that it's comparing the thing it's near to all the

remote ones. So if it's tucked away on the edge near a root server, it's going to have to reach out to all the other root servers to compare against the local.

BRAD VERD: Well, what we've proposed for correctness so far is just a simple DNSSEC validation. That's something that the probe can do on its own without – it's not like a consensus-based –

KARL REUSS: Sorry, I was thinking of staleness there.

BRAD VERD: That one is harder, yeah. You're right. That one gets tricky.

KARL REUSS: Yeah. But it also seems that if on the correctness, if we're worried about something intervening, catching us in the middle, if it's not correct, it's going to throw off all the other measurements as well, so that would probably be invalid. So I'm not sure if that's worth separating them out.

In my mind, I picture this as just one set of medium-reach probes. I think where we place our root servers on the network is important, and if we decide to top them all off on the edge corner,

that's maybe not the smartest deployment move and that should be held against us in some set of metrics.

I'm happy with exchange points, and I think if there's some large case of traffic being intercepted, it's going to affect most of the roots and not just one of them.

KENNETH RENARD:

Just one more comment. The correctness at far reaches where packets are being intercepted, that'll be interesting to collect and find out how widespread the problem is, and we see that a little bit with Atlas, but be interesting to report on.

BRAD VERD:

Yeah, I agree. I think that that's where my mind is as well, that for some of these, if there's two sets of probes, you do them at both. But maybe only one of them would be used for the SLEs, SLAs. The others would be used for informational purposes or archival data or whatever, but only one would be for the SLEs.

But I do want to maybe be conscious of the time, but I do want to go back to something I heard a couple folks say, which is that different operators may have different nearness locations or some may be able to put them closer than others, and that should be documented. And I guess to me that sounds a little bit like a

scaling problem, like you don't want to have – that's a lot to keep track of, especially if you're trying to automate this and whatnot.

And sort of what I'm thinking is that to the extent that we're talking about SLEs here, this should be written in a way that if you were not an operator today and if you were applying to become an operator, you would read this and say, "Is this something that I can meet?" If one of the requirements is that you have to put probes in your network in order to measure your servers, then this should be written in a way that you would understand those requirements, they can be in the same network or same rack or whatever, rather than having to say, "Oh, for G root, it's out here and so that subtracts 3 milliseconds of latency because it's a little bit farther away" or whatever. You see what I'm saying?

Everything should be written so that everyone's on the same level playing field, but you can still accommodate it.

OZAN SAHIN:

Duane, Paul Hoffman has his hand up in the Zoom room. Go ahead, Paul.

PAUL HOFFMAN:

Hi. So I've had comments on many of the things going by in the last 15 minutes. One of my biggest concerns is that as we're discussing near and far, we haven't really nailed down which of

the four measurements are met for being near and far. It sounds like there's general agreement that correctness and staleness would be measured near. Latency would be measured far. Availability is one that people say, oh, that's going to be difficult. But I think that the crux of the issue of near versus far comes down to availability.

Since we are measuring basically an individual instance for an RSO, having a probe near just one of the instances would in fact tell you about availability. Putting one far away would not necessarily in a situation where there is not just congestion but blocking traffic in-between, say, all of the far probes.

And then this comes really to the question of how far. So that was one issue. Second issue is how far is too far for latency. For example, some root server operators currently, and certainly some envisioned for the future – and I'm assuming, by the way, that these metrics are not just for the current root server operators, but for allowing new root server operators in.

Some root server operators might in fact choose using RSSAC 042 to put all of their instances somewhat local for a reason. It might be a government who wants to have more root server availability within their country or within their small region. And therefore, far probes would always give them a negative scoring on latency, but also possibly on availability. Thank you.

BRAD VERD:

Thank you, Paul. I think those are important questions. If you have any suggestions, we'd welcome those as well. I forget, did someone else have their hand up in the room? Okay.

And again, getting back to something, Paul, you were just talking about, which types of measurements to make from which sets of probes, I think you're right, you had a proposed set of doing correctness and freshness/staleness from your probes. I guess availability was a question mark, latency from far probes.

And then in addition, I think for informational or research purposes, we would also want to do some of those correctness – at least correctness but maybe the freshness measurements from the far probes as a measure of the level of interference or whatever is going on.

RUSSS MUNDY:

One thing, Duane, is it sounds like we're getting close to agreement for the need for a near and far structure to the probes. Would it be useful to just take a quick ask around the room if we've reached consensus that there's a need to have both types of probes? And then we can move forward for that. or are we too early for that?

DANIEL MIGAULT: So maybe I'm slow, but I'm wondering if we're not making things very complex by having two kinds of probes. I'm still wondering. I'm not sure we're taking the right path, but I'm happy to be wrong.

RUSSS MUNDY: Go ahead, Ryan.

FRED BAKER: I think there is an argument for having two kinds of probes, because you've got two kinds of questions. And this comes back to a comment that I put in the first sentence of the document, and I think the work party needs to resolve the comment.

But the statement in the first sentence of the document is that we're basically measuring SLAs. And for SLA performance, we really only need what we're calling a near probe.

But there are other questions, for example we have a work party right now that is looking at service locality, and do we actually have servers in the right places and that kind of thing.

You're not going to be able to determine that based on probes that are sitting next to the thing under test. And in a larger sense, we have a research question; what is the user experience of the

system? And you're not going to be able to determine that except by looking at it from the perspective of the user.

So to my mind, there is a – I'm going to say need. I would like to say requirement. I think it becomes almost a requirement that you also measure from someplace that isn't sitting right next to the unit under test.

DANIEL MIGAULT.

Yeah, that is true. I think the need for having close probes is really to be able to identify where the problem is or that the problem is not, in our case, at the RSO level, because if we're far, it could be someone in the middle.

So if we want to be able to say that, to assert that – the RSO or the RSS is meeting some SLA, we have little choice than having only probes that are very close.

If we don't have those, the only thing we can say is when things go well, we can say we're meeting the SLA. But when we – we don't measure that we're meeting those SLA, if we're pretty far, we would not be able to say that's the fault of the RSO. So that's the thing.

It's not as simple as I would have liked it to be, but maybe we have no choice.

RUSSS MUNDY: Ryan, Karl, and Wes.

RYAN STEPHENSON: My turn? Okay, great. Thanks. Russ, I would actually like to talk to you about some type of user experience monitoring, maybe for the root server system, kind of something where – this is slightly off topic from near and far, but I'll get back to that – the monitoring function in whatever model is conceived maybe has some type of dashboard that can be viewed, much like a user experience monitoring to say, “Oh, this is kind of from this area, this area,” maybe geographic area or maybe from – and this is going to tie back to the near, the Internet exchange points.

I understand the difference for near and far, and I think that's great. Personally, I think near should be at these major Internet exchange points, and that's what I'm going to call them. And there may be different terminology, but high peering availability points.

The other thing for the far, I think far would be just anything that's beyond that, whether it's one hop, two, as Ken was talking about, 15 hops. I think there should be, of course, a mixture, but anything that's outside of those – and I really like Brad's idea of pick ten and then each year we – or not each year we, but

whatever function that does this with reevaluates those ten to say whether they're still the correct ten or the ten have changed.

And it doesn't have to be ten. It can be 12, it can be 15, whatsoever. But the far, I would say, would probably just be anything that's beyond that point. So I think maybe we need some definition of near and far if we're going to go down that route, but I also think, again, maybe some type of user experience monitoring dashboard would be kind of – well, first off, it would be neat. But at the second time, I think it would be quite useful just for general people that want to take a look at the root server system, its general overall health, and maybe even break it down into each root operator for these particular locations.

And with the caveat that each root operator doesn't have the same architecture, each root operator has different geographical diversity, which could of course affect some of those measurements, but that's fine, we expect that. That's part of the diverseness of that root server system. That's all I had. Thanks.

RUSSS MUNDY: Thanks, Ryan. I think Karl was next, and then Paul.

KARL REUS: Yeah, I was going to follow up and agree with Daniel that it feels like we may be moving down a path of complexity with near and

far, and we've got different definitions of near, near being near near, and I think I'd agree with Ryan that locating an exchange point seems to be a good near in my mind.

A question, if we did have probes run right next to the root instances, would those be run by third parties, or would they be run by the root operator? Because that would affect where I might need to place them.

BRAD VERD: Well, I think that's up for discussion. Brad was talking a lot about self-reporting, so that would be run by the operator in that case.

RUSSS MUNDY: Paul, Ozan.

PAUL HOFFMAN: Hi. So this discussion from Ryan about user experience and such brings up sort of another topic which I think is central but is going to maybe make a lot of people scream. The root server server selection that's being run by all of the resolvers, which is really what we care about, this is not a user experience, it's a resolver experience, that's going to quickly ignore servers that are far from it.

So a resolver comes up, it probes everything that it's got. It maybe sends one query to each letter, or something, but it very quickly stops even trying to talk to anything that is not near it.

So availability and latency from that resolver to the ones that it's ignoring are completely irrelevant. There is no user experience for that. So as we're saying, "Oh, well, there'll be measurements for availability and measurements for latency," and these are supposed to reflect something for the Internet, that's not the way that we want people using the root server system. And by and large, at least from looking at data that I've seen from a number of the root servers who report data, that's exactly correct. That is, they quickly ignore you if you're not the closest one.

So I'm not sure why we are thinking that we can put labels of things like user experience, or even resolver experience, across all the RSOs if for a particular resolver, it's not going to care about those. Thanks.

RUSSS MUNDY:

Thanks, Paul. Yeah, I agree, that's an important point. And one of the things that we have on our list to discuss later on in the metrics work party session is whether or not the probes themselves – currently, the document talks about making use of a local resolver to actually do the resolution.

One of the things that Duane and I have discussed is it might actually produce more valid results of what we're trying to get measurements of if the probes must do the resolution themselves rather than using any local resolver. But that's a discussion for later on. I think it's a good point you raise, because resolvers are trained and are programmed to ignore stuff that's slow. So I think Wes is next.

WES HARDAKER:

Thank you. It seems to me like we're going to go back and forth about this forever, because we're trying to define it at somewhat of a very high level of whether we need these things, and so my thought was we have a list of metrics that we want to measure, and whether we can use near or far probes depends entirely on basically what's being measured and with what accuracy and/or error injection that you get from being taken place at the right or the wrong place.

So my proposal was that we table it until we go through the document and say the right thing to do is look at each of those measurements and itemize each one whether it can be measured from a near point or a far point, and what the ramifications are if you're measuring it at either one of those.

So if we actually just add to the table of attributes, this can be measured at near or it can be measured at far, or if you do

measure it at far, then there's this issue. I think it would allow us a much better focus afterwards to go, “Oh, no, we actually need both because it turns out this one can only be done near and this one can only be done far according to our objectives.”

RUSSS MUNDY:

Okay. Thanks, Wes. I think Daniel was next.

DANIEL MIGAULT:

One thing I have some hard time to parse is that the probe that is near one RSO might not be near the other RSOs. So each RSO might have its own near probes. But Duane mentioned as well as Brad that self-reporting would be something fine. So in that case, I think it could be – I thought that the metrics were only being done by a third party, so outside the RSOs.

So I figured out that we needed to set the specific platform for that. And I think what I was struggling is that, do we have to replicate something that each RSO already have internally? But if we have self-reporting, then I would say I'm fine. I think we're not duplicating the work. And if we sort of agree that self-reporting would be fine, I think RSO would have the best local metrics to base local numbers to provide the most accurate ones. So it might be a good way to take that path with self-reporting.

RUSS MUNDY: Yeah. Thanks, Daniel. I don't know that we've actually defined yet the entity that will be responsible for doing the measurements, so that's still a topic that is open for discussion, self-measuring and other separate party or something, some mix of some sort, an existing party or something. Liman, I think you were next.

LARS-JOHAN LIMAN: I have two comments. One specifically to that is that I've always envisaged the measurement to be performed by the performance measurement function or subcontractor thereof. I will agree that if we need to measure things within the system, the root server operator will have to be engaged in that and provide data from internal systems, i.e. self-reporting. But for anything external, that should be operated by the [inaudible] function or something falling out of there.

And as a comment to Paul, you mentioned the resolver behavior and ruling out servers that don't respond. I agree that that's how it seems to work in most cases, but we still need to measure because if we have one root server operator which has a latency of one second for every single query, they're not contributing to the system.

So we still need to measure the performance of the root server operators, and in my view, there's this big difference between measuring individual server operators or instances if you wish

and measuring the system as a whole. So to me, it's quite okay if you at one point in the network have a couple of root server operators that have one second latency but there are three or four others with just five milliseconds. Then that's a fine situation for that local point in the network.

But if all the root server operators have a one-second delay, we have totally different situation and that's something that might need to be addressed depending on the impact for that local community and what have you.

If we're talking about the local community of the moon and all the two people there suffer from a one-second delay, I might be willing to live with that. Thanks.

BRAD VERD:

So I wanted to jump in, I guess. When I was thinking about near – again, getting back to the idea that if you're being held to an SLA, you want those as near as possible. I was thinking that a near probe would be like in the same rack as the server. But I heard a couple people say that they would be okay with a “near” probe being located at a top exchange point. And if a lot of people are comfortable with that, then I would say that maybe we can do both from those points. Maybe there's no distinction between near and far.

So that’s something that I would like to explore further, I guess. If the operators are willing to be held to SLAs from measurements made at exchange points, then maybe we don’t need two sets of probes.

RUSSS MUNDY: Go ahead, Ryan.

RYAN STEPHENSON: So to play devil’s advocate on this, say you have a root server operator that’s not performing very well, but knows that if they can – and maybe if they can get a root server instance and that – I hope I’m using the right terms – in that top ten Internet exchange point, that’s immediately going to boost their performance quite a bit.

BRAD VERD: And it’s also going to benefit the Internet.

RYAN STEPHENSON: It’s going to benefit the Internet. True. But then we don’t want these ten locations, or 15 locations, 12 locations, eight locations and so forth being the prime points that operators decide, “Well, I really want to look good, I’m just going to throw my equipment here” and such.

BRAD VERD: I don't know. If you're at the top end exchange points, then you should have pretty good service for the whole Internet, I would think. That's kind of the definition of a top exchange point.

RYAN STEPHENSON: Very true. So I just wanted to maybe kind of put that thought as just an earworm. By the way, thank you, Paul, about the UEM or user experience monitoring with the recursive resolver.

BRAD VERD: Wes, go ahead.

WES HARDAKER: Yeah. The thing that worries me the most about designing this whole system, if we only select a small number of things and we select the top ten exchange points, it's too easy to game the system. So we have to be sure that the measurements that we are taking are actually reflective of the Internet, not of just the measurement system. So I could set up a Raspberry Pie with BGP, announce it only to the local [inaudible] that's actually doing the measurement, and I'm going to look fantastic.

And the other thing that's going to happen is all 13 existing addresses are going to get deployed to those ten exchange

points, not because it's the right thing for the Internet but because it's the right thing for the measurement system. And I don't necessarily have a good solution around that, but I'd be really hesitant to design anything that specifically allows that gaming to be possible.

BRAD VERD: I hear what you're saying. I'm not sure that it's such a concern. We should talk more.

OZAN SAHIN: I have Paul, and then Terry in the queue in the Zoom room. Paul?

PAUL HOFFMAN: Hi. Duane, you just said something that – so I have not been – since I'm a caucus member, not an RSSAC member, I wasn't involved in the development of 037, but you just said something that I think also a little bit goes to the crux of using these measurements for 037, which is 037 is supposed to be looking at current root operators and potential future root operators. And Duane, you said if the operators are willing to accept measurement from the top ten, that could inherently only be the current root operators.

So saying the current root operators are willing to take a certain kind of measurement system could easily prevent future root operators who wanted to do something different architecturally or again locally and such, so if this group is comfortable with just asking the current root operators – well, we know who they are, they're in the room, but I felt that 037 was actually sort of leaning towards the idea of the same sorts of measurements that we would use to measure the current ones would also be used to measure someone coming in or once number 14 through number whatever are in to make sure that they are meeting sort of the same requirements. Thanks.

RUSSS MUNDY:

I'd like to jump in and respond specifically to that. I think you're correct, Paul, and there's actually another element of 037 that – I don't know if people haven't noticed it that much, but I think it's called the architecture function where 037 talks about having measurements to be able to assess the overall need of the root server system if it is sufficient overall or if more or less RSOs are needed.

So it's both the combination of each of the RSOs would be related to the metrics system, but the overall measurement of the overall system needs to be comprehensible in some manner from what we develop here.

So it's current RSOs, future RSOs, and whatever the RSS consists of at any particular time. That's what we're trying to get a good set of metrics for here.

OZAN SAHIN: Over to you, Terry.

TERRY MANDERSON: Thank you. Terry Manderson from ICANN Managed Root Server. The 11 principles that we undertook within RSSAC is important to me. One of those principles was diversity of operation. If we construct measurements that entice RSOs to be on the same fabric, such as an IXP or such as an exchange point, we're really leading into a sub-optimal situation where they're all going to share faith. I don't believe that's particularly healthy for the Internet.

Perhaps even so, an RSP chooses not to go into an IXP, and then they're disadvantaged in the metrics measurement construct. So I'd like you to keep that in mind. Thank you.

BRAD VERD: So Terry, would you then advocate for near probes being close to RSOs? I.e. in the same rack or same ISP.

TERRY MANDERSON: Perhaps in the same rack, provided there is adequate fabric separation between different or individual RSOs.

BRAD VERD: Thank you.

TERRY MANDERSON: Does that make sense?

KARL REUS: One thing I've not heard us discuss, and maybe it took place before I joined RSSAC, is do we want to allow for the concept of just [a] geographically local root to a particular port of the globe, a particular country or area? Or should all roots serve the globe equally?

I don't know whether we've actually answered that or not. I've been assuming that all roots should serve the entire Internet community and not just the local region. So my thought for measuring from exchange points would not just be to pick the top ten, because I think that serves the current folks running, but to make it a mix of top ten plus a certain number of probes per continent, something along those lines so that we do get coverage to less served areas.

BRAD VERD:

Right. I'm struggling to respond to Karl. So I think the question that I would like to get clarity on is, do we need two sets or not? If we just need one set, we can put them at the top ten, the top 100, the bottom ten, whatever. My question is – sorry, Paul, but I can only ask the operators who are in the room who've expressed discomfort at being held to SLAs, are you okay with being held to SLAs from exchange points, or not?

UNIDENTIFIED MALE:

There's kind of a Heisenberg issue of where the points are. I apologize for missing the first half hour, but were we going to develop a series of measurement points to measure this or try to use some existing one? And I think what numbers we try to hit will depend on those.

Real life example, yesterday we were approached by the government of the most populous nation on Earth to put root server instances in the hinterlands where previously they'd all been in the capital. And I was thrilled, and I would hate to think I'd have to say, "No, we need to put them near those racks where they can help way fewer people because of the measurements." That seems like a silly motivation for us to be adding.

Rather, I would think the nation of – arbitrarily, let's pick China, it'd be interesting to see how you can get resolved from the Uyghur homeland or an obscure part of Nepal, or something else.

I'm almost wondering whether being in the population centers is doing the wrong kind of measurement.

BRAD VERD: So, what do you propose? Do you want 15 darts and a Mercator projection map?

UNIDENTIFIED MALE: [inaudible].

BRAD VERD: For the probe locations?

UNIDENTIFIED MALE: Yeah.

BRAD VERD: Okay, got you.

RUSSS MUNDY: Well, the question that we're really trying to get our arms around at this point is, does this work party think it's appropriate to have two – I'll call them distinguishable set of probes, one that would in theory be located topologically close to the RSO entity being

measured and some set that would be at some topological distance from that?

Or if as perhaps Wes suggested earlier, that we look at each of the metrics we're talking about and figure out what makes sense for that metric, and then go back and ask the question, does just saying a single type of probe meeting whatever the requirements are for each of the metrics make more sense than – because we've been trying to get to the question of there's a lot of concern about SLAs and being held responsible in an SLA for things you can't control. And that was where the idea of near and far probe evolved from. So that's where that came from, that's why we're trying to get that question resolved.

UNIDENTIFIED MALE: At the risk of jumping on really briefly, I just want to clarify that. By being close, you're sort of in the catchment, so you can guarantee what thing it is you're measuring –

BRAD VERD: [Maybe in the same rack.]

UNIDENTIFIED MALE: Okay, but it's kind of artificial. Where the real measure is, Joe Random somewhere who's within a two-day drive of ten

different instances is a much more realistic example of they're going to get resolved by somebody somewhere.

RUSSS MUNDY:

But for instance, it's not unrealistic for correctness. So that was one of the reasons why I think Wes brought up earlier, let's look at each of the metrics to see what makes sense there. And I think – have you been watching the [order], Duane?

LARS-JOHAN LIMAN:

Can I ask that you maintain a proper list of speakers? Because I think this is jumping back and forth a little too much. So when we come to measure from the ten [IEXes] we could do something like we select not ten but we select maybe 50 or 100, and we say you have to be near to a fraction of that, you have to meet that nearness criterion from 10 out of 100.

And we also need to look at, again, individuals versus the system. So the system has to be reachable, let's just say three roots have to be reachable within the limits from all 100. But which constellation of three may vary between the [IEXes.] That will give a better image and possibly also better deployment as a whole of the whole entire system.

When it comes to geography, again, we need to measure the system, not individual roots. The guy two days' drive off, is he

okay because he can reach two servers but not the entire set of 13? I would say yes. If he cannot reach anyone, maybe you have a problem, but again, he might be an outlier.

So I've been thinking about the measurements as a mass measurement where you have a whole lot of probes. Think of it as the Atlas system where you measure both the system as a whole and the individual server operators and where you cut off the outliers.

It's okay if 5% have a latency of more than X, because you cut off the trailing ends and you look at the majority, the percentiles again. And that's where you need to be responsive and make sure that you fill the needs. That's my view or my way of thinking about measuring the system.

For closed ones, I think we probably need to have closed ones for our own benefit. I as a root server operator like to watch my servers to make sure that they perform locally, so I'm sure that's not the problem. But I also want to know if Jon Doe in the forest is able to reach me, and I cannot measure that myself. I need help from a probe system, one with lots of probes spread around the globe, and preferably even different systems that measure from different angles.

DANIEL MIGAULT:

I'd like one clarification. When we say the measurement can be taken from the same rack, does it mean the RSO is providing this measurement, or someone can go into that rack? It's only the RSO? Okay.

So it will be self-measurement in that case when we say the same rack. Okay.

The other thing, back to what Liman just said, I think it's important that when we have the measurements, because we are trying to – there will be a need at some point that we're able to consider a subset of the different probes, and we're saying maybe a subset – providing the measurement from a subset depending on the geographic locations, depending on the location [to and next to each client,] we're defining some different subset that can provide measurements that can be interpreted in a different way.

So I would go back to – I believe it's important that we're – for a further analysis, that we know which probe and we've got information on each probing point so that we can establish which is the right subset. And it could be far, near exchange point, other things. But I would say it might be for further studies.

It seems from this discussion that it might be also good that we don't really care if it's far or near as long as we have a third-party platform and self-assessment, self-reporting. And that would

match maybe the near and far aspect, and it doesn't need to rebuild two times the same platform in the case of the near one.

So I think it's good to have this self-reporting and another platform.

RYAN STEPHENSON: I'm going to defer to Wes before I speak.

WES HARDAKER: Go ahead, I'll go after you.

RYAN STEPHENSON. No, I actually want to hear what you had to say.

WES HARDAKER: Alright. Thank you for yielding your time. The rebuttal will come in a minute, I guess.

BRAD VERD: By the way, we're up against the break, so go ahead.

WES HARDAKER: Okay, so I will go quickly. I mentioned this point in the workshop, that one of the ways out of this situation of designing for a system which is perfect in ten locations is that the SLAs offered to each or

held to each operator may differ, because if I have a root server that I want to specialize, helps the people in remote places, like deepest, darkest [Peru so Paddington Bear] can get good root service, then they won't be held to an SLA that says you must be latent within 100 milliseconds, because the reality is that the measurement system is not going to be able to measure that particular goal properly.

So in the same way I also mentioned like a nuclear bunker, so that one would survive for the longest period of time because they're spending all their infrastructure costs on concrete, not on diversity.

And the reality is that a particular threshold, we might need to have different threshold values and SLAs. We've written some examples in the document, but it may be that a particular proposal responding to an RFP for doing something new is actually going to say, "No, this doesn't work for me, I want this one instead. I want a faster one because I'm going to be the least latent server out there."

Who knows what it's going to be, but we shouldn't hold everybody to the same standard, because the goals and the diversity needs will differ.

RYAN STEPHENSON: So I'm just going to talk specifically in my root for right now, and I'll make this quick. G root, we pretty much cannot hook up a probe, so it would have to be – if there is something for near, it would have to be defined in a geographic area as far as near.

The other thing is I kind of – I'd like to retract a lot of things that I said about, hey, the ten points, great, because I'm thinking more of the terms of what Liman and Wes have said. It's entirely sure to game the system, and there has to be preventative measures to prevent the gaming of the system, which I'm wondering if then this near and far discussion is kind of moot because you have to take a look at measurements – you do have to craft, as Wes was saying, and to reiterate Wes, you're going to have to pretty much craft the SLAs or these metrics towards – you have to think of the complete diversity of the root server operators themselves, where they're located, where they're held, so forth, their network topologies, just because of the fact some are scattered across the earth, some are smaller in scale, but handle the global instances quite fine.

So anyways, that's it.

BRAD VERD: Alright. Fred, you're the chair. We're overtime. What do you want to do?

FRED BAKER:

Well, I'm just thinking I'd like to get a comment in that I actually raised my hand about two hours ago, and that is that I think it's possible to [rathole] on use cases. One of the use cases for these kinds of metrics that I see is that almost every ICANN meeting, I am approached by somebody that says, "I would like to have a server instance in my country as opposed to the one next door," whoever that is.

And it seems like there's got to be a reasonable case of saying, well, fine, take one of these measurement systems and tell me what's different between your country and the country next door, and allow them to make an informed decision as opposed to a protocol decision or whatever was driving them.

So I think far needs to have enough possibility of variation in definition for people to apply it in ways other than the very narrow way we're talking about it this morning.

Having said that, you're the chair of this meeting.

UNIDENTIFIED MALE:

Break. [inaudible] 10:30?

BRAD VERD:

Okay, have a ten-minute break.

RUSSS MUNDY: We have coffee over there, have juice, and I think some water. Please be back in your chairs by 10:30. Keep the good discussion going.

[END OF TRANSCRIPTION]