DUANE WESSELS: I want to share some data and some graphs that I had worked on for a separate project that are kind of related to this, and then after that, we're going to go back into looking at the document that was up yesterday and there was a bunch of TBD items, and we're hoping to have some discussion about where the people in the room and on the phone, what they think some of those values might actually – what might be good values for some of those TBD items.

Can you put up the first graphic for me? And the remote folks can see it as well, right? Okay. So I've got some graphs to share. This is data that is actually from RIPE Atlas measurements, and a couple things about RIPE Atlas if you don't know. First, there's quite a few of them, something like 10,000 or more now, and every RIPE Atlas probe, it's a little box, fits in your hand kind of thing, and it's designed to do measurements all the time.

One of the measurements that every RIPE Atlas probe does is send a DNS query to all of the root nameservers, both v4 and 6, I believe at an interval of every four minutes or something like that, so relatively frequently. This data is publicly available, so last year, I did a project where I downloaded a lot of this data and ended up making graphs that look like this.

This is data for one of the root servers, anonymous, and what it's showing here is for every day, you can see the X axis on the bottom goes from 2012 to the end of 2018, I think. So for every day, there is a five-minute window selected and all of the measurements for that

window are retrieved, so about 10,000 measurements, then take the median value and plot that on this line.

This particular root server, its median latency as seen by RIPE Atlas, is generally in the range of 50 milliseconds. There are certainly some spikes up higher, and there's a little bit of a difference between IPv4 and IPv6, but you could say that it's usually in the range of 50 milliseconds.

There's this slight downward decreasing trend in this data, and my theory, my intuition on this is that that is because the number of RIPE Atlas probes increases over time, it's not because of anything that the root server operator has done to change or improve their service. It's really because there are more data points from more RIPE Atlas probes. And you'll see that that's a common trend in all of these graphs.

This is six years on the X axis. I think that's about as far back as these RIPE Atlas measurements go. It's really a huge amount of data that they have, and this is one measurement per probe per day. You could really get one measurement per probe per five minutes if you had the resources to download and process it all. So it's a great resource. Alright, let's go to the next one. So that's one.

This is another one. Obviously, different operator, and in this case, the median latency as measured by RIPE Atlas is in range of about 120 milliseconds pretty consistently across the board with some spikes as seen in the other one. Let's go to the next one, please.

This one is lower. We're, again, down below the 50-millisecond range, and again, you see the downward trend, which hi assert is from the increasing number of RIPE Atlas probes over time.

UNIDENTIFIED MALE: [Decrease in average distance.]

DUANE WESSELS: Right. Yeah. So this is median values, but exactly right. Alright, let's look at the last one, please. So here's one where you can see there have been some sort of dramatic changes over time. It started out in that 12-millisecond range and then something happened and it dropped right down, and then it looks like the green is IPv4, so for v6, it went back up a little bit for some reason.

So I just wanted to show this to give people a kind of sense of what kind of latency measurements you can do with a measurement platform and what results might look like, and sort of a sense of where the ranges of, say, a median latency would like for a range of operators. Again, this is median. Yesterday, we talked about maybe doing something like 5th percentile, which would drop that down even lower. We could do 95th percentile if that was interesting to you, and that would be up higher. But just to give people a sense of kind of what's possible.

So those are four individual operators, and then one more slide, Steve. This is different, and this is something that RIPE Atlas added more recently so that here, the X axis range only goes from I guess the start of 2017 to the end of 2018, two years. This measurement, the probe sends a DNS query for a randomly selected TLD that would not be in the root zone. So this is designed to be a cache miss sent to the probe's local recursive nameserver. And this is the latency of that. On the Y axis, this line is in the range of 30 milliseconds.

And this measurement was created actually in response to, I think, something we were discussing either in RSSAC or Root Ops, I forget, but the idea was to have a measurement of the root server system, not individual operators. So that's what this is, and it seems to work pretty well. Any questions about this data or these graphs before we move on?

UNIDENTIFIED MALE: [Chrome does this for us, doesn't it?]

DUANE WESSELS: Chrome does something similar for different reasons, yes. It does send probe queries very much like that.

UNIDENTIFIED MALE: [So we could tickle the Chrome people to give us the numbers.]

DUANE WESSELS: Yeah, perhaps. I don't think they collect that data, but anyway –

UNIDENTIFIED MALE: [They could.]

DUANE WESSELS: They could, yeah.

RUSS MUNDY: One of the things that is actually not on our list of things that are unknowns or two be discussed, but Duane, Steve and I came up with it last night as we were talking about today's, is later on in the discussion, because it really applies to all the metrics, we might want to discuss, do people have a sense of how many probes or test points make sense? Atlas is perhaps the largest at least widely recognized one. There may be other big systems. I'm not familiar with them. But 10,000 is a lot of probes.

But Chrome, how many installations are there that are doing things of this nature? Yes, Brad.

BRAD VERD: I don't think you need a platform like Chrome or Atlas to get something reasonable. I think, Matt, what is ThousandEyes? 60, 70?

MATT: How many probes do they have?

BRAD VERD: Yeah, roughly.

MATT: 3-400.

BRAD VERD: Okay, so that's –

MATT: [inaudible]. There are platforms out there that support that.

BRAD VERD: Yeah. So I think Atlas is an amazing platform, and obviously it's there, if we can use it, we should use it. I'm just saying that we certainly don't need every Chrome instance or whatnot to get some reasonable number that is believable or actionable or whatever you want to call it.

PAUL: I'd like to second that. I also want to point out that Chrome and Atlas are often installed in a place that can't make UDP queries at all, or if they can, they can't be on Port 53 unless they're going to a local nameserver. So it's actually the wrong population. We need something that represents the recursive nameservers placement in the topology and with respect to local policy and firewalls, and browsers and Atlas probes are not that.

UNIDENTIFIED MALE: Paul is correct. I just want to add, Atlas, we have a subnetwork [inaudible] called Atlas Anchors, and they're installed in datacenters, normally without any filter or anything. There are about 500 of them in, I think, 155 countries, so they're well-spread. And yeah, they're normally with service provider datacenters. So they are more stable targets and they don't have most of the issues. So you can actually – and all of these measurements are done from them as well, so you can actually only choose those probes or anchors basically as your

measurement points. And they also do mesh measurements, so they have traceroute and ping every five minutes to each other. All of them. So there's a lot of data there as well. Just between those anchors.

UNIDENTIFIED MALE: I mostly just wonder, are we proposing like building out a probe network, or are we looking at the ones that exist and trying to figure out what the hell to do with it? Or does the existence of the anchor probes settle any of Paul's concerns about using Atlas?

DUANE WESSELS: So in our materials so far, we are saying that measurements would be taken from some kind of distributed platform. Whether that's something we have to build or not, we don't know. I would like to have input on whether – let me think about this. I think we need to decide if we want to design measurements such that they can be run on existing platforms or not, or do we want to make them so that they run on the platform that we build? So if we design measuremetns that can run on an Atlas or a ThousandEyes or what have you, then we have saved a lot of work and money for ourselves, probably, right?

But if they don't have the features or whatever that we need, then we may have to go and design or have someone build something specific to the root server system metrics.

UNIDENTIFIED MALE: [inaudible].

DUANE WESSELS: Yeah, and that's something that I would welcome discussion on, what people think.

LARS-JOHAN LIMAN: I would appreciate definitions of tests that work on multiple platforms, that are generic. But when we design them, we should take care so that we don't force ourselves to design our own probe network. So we should look at existing platforms and see if there are commonalities, see if there are specific differences that we need to navigate around. And we have a few known platforms, ThousandEyes, we have Atlas, Atlas anchors and so on.

And I think it's actually a fairly easy task to just navigate around known problems if we see them, because that way, we will be able to compare results from different platforms, and that could be interesting in the future.

DUANE WESSELS: Okay. Thanks. Steve, can we go back to the Google document screen? So on the screen, this is the sort of strawman latency metric, and a list of parameters that need to be agreed on and decided for example. The other proposed metrics, not surprisingly, are sort of similar, and so for example the measurement interval, it would make sense if you decide it should be ten minutes for latency, then it should also be ten minutes for correctness and the other things. So these may be the kinds of things we only need to decide once, but they're on this screen here, and on

behalf of the work party, I would welcome comments and input on what some of these values should actually be. So, how often should a probe do its measurements to a root server? Should it be every minute, should it be every hour? What's a reasonable value? Keeping in mind that the more often you measure, that's more data to manage and do calculations on and things like that. So let's maybe start with measurement interval.

LARS-JOHAN LIMAN: I think the question I ask myself when I try to get a feeling for this is, how much more information do I gain from querying every minute as opposed to query every five minutes? What's better with more often there? What's better with five minutes compared to an hour? That's where I start to feel like there's probably a difference there. Five minutes, one minute, probably not. So I can home in on five minutes as something that – there are changes that are short interval changes, and they can probably fix themselves in a few minutes, but if it's more than five minutes, then you may want to have that indicated somehow in a test. So that's just my gut feeling.

DUANE WESSELS: Okay.

UNIDENTIFIED MALE: So my thoughts are, depends, if you only care about latency, I don't think a very frequent interval is needed. One minute versus five minutes is not going to make a difference when we're talking about weeks,

months, years. If you care about availability, then you certainly want to be probing more often. And whether these measurements also dovetail into an availability check is another story as well. So if my thought is if it's availability, it should be a short interval, short being five minutes or less. And if it's latency, it should be, honestly, I think over time I think 15 minutes is generous.

DUANE WESSELS: So I do think, Matt, that – sorry, I forgot about you. Paul, go ahead.

PAUL: Sorry, I just entered that I lowered my hand because Matt just said what I wanted, which is that in fact, these measurements are not all parallel. And remember, we still don't know why we're taking these measures. So I think it's a bit premature for us to be picking times for each one, but I fully agree with Matt that availability is something that you measure more quickly than latency or correctness.

WES HARDAKER: The other thing to take into account is when it comes to monitoring, you really want to alert on the first one, so if you get a high-latency query and people are actually watching and want to make decisions off of it, you should really wait for three or four or something like that in order to make sure it wasn't some weird network glitch, so you may need to increase the frequency just to deal with that.

UNIDENTIFIED MALE: We're not talking about alerting here though, right? We're just talking about gathering the measurements. Alerting would be probably something separate.

WES HARDAKER: Yeah, but how frequently you gather depends on what you want to do with it, as Paul just indicated too.

UNIDENTIFIED MALE: Right.

DUANE WESSELS: Yeah, so there's a related undefined parameter here which I call aggregation interval. You would make measurements at some measurement interval, let's say five minutes, and then you would aggregate those over a larger aggregation interval, maybe one day, and h ten report that as a median or some percentile or something like that. So those two are certainly related. I think you wouldn't want to have them both be five minutes, that would not give you a very good data sample. So they are definitely related.

So I guess I've heard five minutes is probably in the right range-ish for what we're thinking, but what about – let's maybe have a discussion about aggregation interval. So, is daily about the right range for this, or is there a reason to – and again, as you said, this is going to be sort of a long-term collection and may be used by whatever the PMMF is going to call it, so this is going to be used by some future function in

monitoring and looking at the root server system, so how often should those measurements be published, I guess.

So in RSSAC 002, obviously, we do that on a daily basis, and there are certainly – I can speak for myself – times when I've been looking at that RSSAC 002 data and wishing that it was hourly or something like that, but I think for the most part, it's pretty good on a daily basis.

WES HARDAKER: I think for purposes of what those are going to be used for, daily is fine. I would like to add standard deviation to the list as well.

DUANE WESSELS: So certainly, standard deviation could be calculated. That would be a straightforward thing to do. Do you have an idea of how that would be used?

WES HARDAKER: Yeah, because it helps you detect whether the service is stable or not. If you have a suddenly increased median latency with a wide standard deviation, I'm not sure it's a service that's being provided. It's probably network stuff in-between. It gives you just another indication of what the root cause might be.

RUSS MUNDY: Question on that. Standard deviation is a fairly complicated function to perform in a simple thing like a probe. Figuring out the variance would be – standard deviation is square root of the variance –

WES HARDAKER: I don't think aggregation is going to happen on the probe though, I think that'll happen on the server that's collecting the data.

LARS-JOHAN LIMAN: The probe would just report.

UNIDENTIFIED MALE: So a couple people said a day is appropriate for what this is going to be used for, but what is this going to be used for? Do we really have an understanding? Duane, what's your concept in proposing the aggregation interval? Because presumably, all these individual data points are going to be saved, so anybody who wanted to go scrounging through them could get them all and we're not talking about a huge amount of data. So I'm not saying don't aggregate, I'm just saying it's hard to talk about it without knowing what use cases we have in mind.

DUANE WESSELS: Well, I guess the only concrete use case we have is what we have in RSSAC 037, which is there would be some function whose job it would be to perform these measurements and watch them to evaluate the root servers and the root server system for its performance

characteristics. Other than that, I haven't heard any concrete uses proposed, but we've been tasked with defining metrics. Go ahead, Brad.

BRAD VERD: I can think of numerous things this could be used for and probably should be used for.

UNIDENTIFIED MALE: To be clear, I'm not saying it's not useful. I'm just saying in order to have a serious discussion about interval, we'd want to understand the use cases to know.

BRAD VERD: The only public data on the roots is RSSAC 002 and/or other public monitoring systems. Terry is not here, I can't put him on the spot, but ICANN says they're responsible for the global DNS. I've heard that from a couple people based upon the interpretation of the bylaws. It seems like if we were reporting it, it would carry much more weight, be much more transparent rather than this closed group of operators. It seems like if we made it available to the community, to the research people, to whatever, again, it's a transparency thing, and I don't want to build on 002, but I think this is – I'm trying to think of the right word that is not maturing of a platform, of a group, of creating a – sharing what the metrics are that we hold ourselves accountable to, and providing that to the public seems like the right thing to do to me. 37 aside. So I don't want this to be tied to 37. I don't think it should be tied to 37. I think

this is something that we should have done long ago, and we've tried numerous times, and we've failed each time.

UNIDENTIFIED MALE: To be clear, I'm not saying don't do that.

BRAD VERD: No.

UNIDENTIFIED MALE: What I'm saying is – so that's a good use case, so then what we should –

BRAD VERD: I'm providing the use case. That's what I'm saying. I'm not trying to argue, I'm just saying this seems like a very reasonable use case to me. So I think it was Jeff, and then Andrew.

JEFF OSBORN: If we put something like this in place, isn't it sort of inevitable that this becomes the measurement and reporting system against the SLAs that we will have signed to by then? Because if this isn't that system, then we're going to have to have a whole other one.

BRAD VERD: So again, what you put in place could easily be plugged into 37, which is what Duane said, this would be used by the PMMF. But if 37 never happened, it still seems like this is a good thing to do, a maturity thing

to do, something that probably should have happened long ago but didn't, and that's my use case. I could think of others if I need to, just give me a couple minutes.

UNIDENTIFIED MALE: Well, if this is something that in three years one of us is blowing an SLA and is looking back and saying, "Oh, hell, if I'd realized this was going to happen, this is what I would have done back in April, whatever it is, 2019," that's probably worth all of us thinking about, because it seems pretty likely this is going to be a measuring system against which we are measured, and some of this will occasionally fail and we'll wish we'd done something different with it now.

BRAD VERD: While you're correct, my guess is that what'll happen years from now is we'll look back and be like, "Oh, we weren't strong enough on some of those. We need to up our game." Because that's the trend that has happened. We kind of get narrower over time as we – as the system ...

STEVE SHENG: So from Daniel Migault, "Median or average with outliers are always important to characterize the measurements. I believe that should be represented. The representation I'm thinking of is a quartile representation with outliers, median or average. Thanks."

UNIDENTIFIED MALE: Go ahead, Matt.

MATT: I would just support the overall concept that whatever this aggregation interval is – and to be clear, a day seems reasonable to me, I've done all kinds of work over the years where I've aggregated stuff, measurements like this to a day. But I support the idea that both Wes and Daniel is saying, that it shouldn't just be a single aggregation, that our aggregation should result in multiple values, for example like what Daniel suggested. I think that would be really important. And why not throw in standard deviation while we're at it.

BRAD VERD: So if there is a public monitoring system like RIPE, and it is the only public monitoring – I shouldn't say only, but it is one of the only ones, right? It's like the de facto, I guess. If you don't have your own monitoring system, then by default, whatever RIPE says is kind of gospel. So if RIPE says there's an outage or your latency is bad, and you don't have your own monitoring system to prove that it's not, then you're kind of guilty until proven otherwise. So again, I feel like just as a maturity thing, this seems like something that we should have. Not to say that RIPE is wrong or any of that, but two eyes are certainly better than one. But I would certainly like to see if we've answered the question that now I've heard from you, Matt, that was eloquently put, and Paul's comment that there's no purpose for what we're doing here yet. I'm curious if we've defined a purpose yet or if we need to continue and get that answered.

MATT: My implication was never there was no purpose.

BRAD VERD: No. It was [what we're using it for,] and I think that's what Paul said in a different tone. That's all.

MATT: I guess my assumption, really my understanding based on listening to all this is that I'm in agreement with you that this is – I think maturity is a good word for it, it's something that we in this room ought to have our "own" measurements just on general principles. But that 37 is the primary driver for this, because that is going to need measurements for things like SLA calculations and eventual agreements. So I'm in complete agreement with everything you've said about the reason for the measurement.

BRAD VERD: Yeah, I do agree that 37 is the primary driver, but we don't know what's going to happen with 37. 37 might happen, it might not happen. It might happen in a different way than we ever expected. But I believe that this is all stuff that we should be doing and should have done already, we just haven't yet. So I just want to make sure we're all on the same page with that, because I keep hearing, "Why are we doing this?" And it confuses me, that's all.

RUSS MUNDY: If I could add just a little bit, it is an effort by this group of people, at this point this group is the work party, and eventually we'll become RSSAC as it processes through to actually lay out some definition of what ought to be measured, how it ought to be measured, sort of the big level views of it, and I have always said for a long time – I know people have heard me harp on this soapbox that we need to know how to do it as a system in addition to what we're currently measuring [inaudible].

If 37 happens, then it fits nicely. If it doesn't, then we have a description of what and how we're measuring, how the root server system is working, and I think that's the high-level fundamental purpose for the whole work party activities, if that hopefully is helpful. That's at least the model that I've been trying to push for as one of the co-chairs of the work party.

DUANE WESSELS: Okay. Thanks, Russ. It feels to me like the sort of long-ish list of TBDs is maybe a little bit overwhelming for everyone at this point, and so our time might be better spent moving on to some of the other metrics and some of the harder discussions. I'm perfectly willing to take back some of this work to the regular work party meetings and have them make proposals on what some of these values should be, some of the details. There's a lot of details to work out in some of these measurements, for example the particulars of the query and the extent to which [retries] and that kind of thing happen.

Does that say Paul has a hand raised? Okay, go ahead, Paul.

PAUL: Two things. One is, Brad, thank you. I actually now feel like once we write down what just got said, we do know why we're taking these measurements. It's just we haven't seen that yet.

A second minor point but I think that's important is just looking at what we have here for latency. I want to be clear, and I think this is good, some of these measurements will not be, "Do a probe and measure it." This one in fact, if you look at the second paragraph where we say response is not received after however many seconds are considered to be timed out, we may actually have tests that say if there is a timeout, try again. So not just a simple probe but a make sure that this probe is working, and if it's not, try again.

I think those are perfectly reasonable tests. They make the data much harder to interpret, but they take into account the fact that the DNS is full of crap.

DUANE WESSELS: Yeah. I agree, Paul, that we may indeed have to do that kind of – [we won't get] that specific. I guess for me, one of the unknowns still at this point is a lot of these proposed metrics are such that you could imagine sending one query to do multiple types of measurements, for example this query could also include a request for signatures and that kind of thing, so then you could do the correctness check, you could validate that.

So I guess I would suggest that as we go through some of these other metrics, let's keep in mind places where you can have that optimization where it makes sense to try to do two or more different metrics with

one probe for example, and where it makes sense to have the same retry policy for those different types of metrics for example.

RUSS MUNDY: [inaudible] being recorded, so we can always go back and listen if we don't get captured right now.

DUANE WESSELS: Alright, so let's pick on that a little bit more. So we were just talking about retries. Anyone in the room or on the phone have any thoughts on – if you're probing, you're doing a latency measurement, how persistent should you be? Should you retry three times, four times, mimicking the way a recursive nameserver would retry Or keep things simple and just do one and say, well, if it's gone then it's some other problem and I'm not going to use that in my latency measurement?

UNIDENTIFIED MALE: My opinion is try once. If you keep on retrying, that's obviously going to impact– then you're going to have to worry about if it works on the second retry and how do you calculate the latency and how that measures against a different server where it always responded on the first query? So my thought is no retries.

DUANE WESSELS: No other opinions, apparently.

UNIDENTIFIED MALE: Plus one.

UNIDENTIFIED MALE: Yeah, plus one. I agree.

LARS-JOHAN LIMAN: I agree with your sentiment. It would be interesting though to know if you have lots of outages. So that may be a number that you want to collect.

DUANE WESSELS: I think that's covered under maybe an availability metric.

LARS-JOHAN LIMAN: Right. Fair enough.

DUANE WESSELS: So the question is if you're only trying to measure latency, how persistent do you need to be to get that latency measurement? And as Matt said, what does it mean? If you retry after four seconds, does that mean your latency was four seconds?

LARS-JOHAN LIMAN: I was thinking about that, but I agree with Matt, so it's a moot question. But you could argue that if you send the first question and it doesn't respond in four seconds, you send a second one which is responded to in two seconds, the actual delay would be six seconds. But we're not

going down that path, I think. But you have to somehow note the fact that there's an outage, and that should influence the number that you report back to the average calculator and so on. And maybe the number you report is four seconds ...

UNIDENTIFIED MALE: Which is why I'm suggesting you don't do a retry, because you have a level playing field. If some of the latency numbers are based on no retries and one particular server you get a lot of retries, it's not a level playing field.

LARS-JOHAN LIMAN: But do you agree that if you have an outage, then you report four seconds with the given example here? Or what do you report?

WES HARDAKER: That's not the job of the latency. That's the job of availability. So latency is measuring the time. And let's play it out for a day, we're aggregating to a day, then yeah, that reporting on that particular metric is broken. You don't report any value, it's nonexistent. No.

UNIDENTIFIED MALE: Wouldn't a recursive server – even if it were to retry, it would retry to a different nameserver in the ns set, right? It wouldn't retry to the same nameserver if it didn't get a response, correct?

UNIDENTIFIED MALE: I'm sorry, I'm going to jump in. No, that's not correct. Different resolver software does different stuff. I'm going to jump in here and say maybe it's useful for us to figure out what – are we trying to measure things based on how the resolvers that our customer are doing things? And if so, we have to look in their code, because they do things – and they change what they do over time.

PAUL VIXIE: Even if it didn't change what it did over time, even if it were standardized, and even if all of the resolvers out there followed that standard – none of which by the way is true – this would still be the wrong approach. In specific answer to Matt's question, most of these resolvers will try all of the addresses belonging to a particular nameserver, in other words the A and AAAA records that are owned by the target of an NS record. They'll try all those before they move on to the next nameserver, the next NS record and go through all of its various addresses. But others do it row-wise rather than column-wise, or randomize it, or they'll do RTT sorting. There's a whole bunch of different ways to do this.

So the typical way that a DNS operator accounts for this just in the enterprise context, and certainly the way that BIND's syslog metrics always worked is that there was not a counter for retry. If you sent a question and it timed out, that was a timeout, and then later, if you sent a question and it was answered, then that was an answer and you would measure its latency. But you would not associate – let's imagine particularly poorly configured zone, only has one NS record with one address. All recursive nameservers that I'm aware of will try that

address more than once. But the right way to think about what it's experiencing is some number of timeouts followed hopefully by success that has a latency. Timeouts don't have a latency. But you would not make a separate counter in any sensible statistics gathering system to say that my retry succeeded even though I had earlier timeouts. Those events are disconnected by the time you're measuring what the thing's doing. Thank you.

DUANE WESSELS: Alright. Thank you very much, Paul, for that. So one other aspect of this metric which is not at all decided is what's called here specifics of the DNS query. That means what name is queried for, what type is queried for, and maybe more importantly, it means what sort of other options you might set in the query. Do you ask for the DNSSEC record so that it's a large response for whatever reason, or choice of EDNs buffer size and things like that.

So for example if it's a large response that might be truncated, then you would have a TCP retry and maybe that's part of the goal of what you want to measure, or maybe it's not, maybe you want to make sure that it's never truncated and so you keep the response small. I guess speaking for myself, if it was my choice, I would probably keep it simple, do a query to the root name itself or some TLD and try to keep the response small, but I'd like to hear other people's opinions on that if they have any.

BRAD VERD: Just representative of a typical query, I think.

UNIDENTIFIED MALE: So a typical query that the root nameservers receive would be like a name with two or three labels that's going to end up in a delegation-type response, is that sort of –

UNIDENTIFIED MALE: Yeah. I guess most of the responses are NX domain.

UNIDENTIFIED MALE: Well, maybe NX domain or maybe not. You could say let's always set the query name to www.example.com, so every time you get a response [it's a] referral to the com zone.

LARS-JOHAN LIMAN: [com].rootservers.org.

UNIDENTIFIED MALE: Or a specific name. Right. Or I guess if you wanted to add complexity, you could say a set of names to choose from or some random component or something like that for whatever reason, but ...

LARS-JOHAN LIMAN: We don't need random components with that. And doing the numbers in my head, I was going to talk about the DNSSEC, [but I would use NSEC] for that. That's not a problem. But the thing we want to measure here is response time, so we should not complicate stuff. I'm actually on

your side. a simple query that has a well-defined answer and short and to the point.

DUANE WESSELS: Okay. Thanks. I guess scroll down just a little bit more, Steve. So based on yesterday's discussion, I think we don't really want to get into the thresholds just yet probably, we want to save that for later. So I'll sort of skip that. But I did also include at the bottom here what an example result might look like. I heard earlier today the discussion saying that something like only median is probably not enough, that there should be quartiles, percentiles, standard deviation, that sort of thing. So that's certainly reasonable, I think. We can add that to the document. Do we have general agreement on the proposal for doing sub-metrics? So for every operator, these latency measurements are done both over v6, v4 and TCP, UDP at this point?

LARS-JOHAN LIMAN: Again, with my warning about measuring over TCP, but yes. Just make sure you know what you're measuring, how you're measuring and describe that.

DUANE WESSELS: Yeah, I think that's a manageable thing. There are a couple of things that I worry about a little bit, like at some point we're probably going to see wider deployment of TCP Fast Open, which could significantly affect latency measurements of this type, but probably wouldn't advise doing that yet. Maybe in the future this would be revised to include that or

something, but I would say it's just a straightforward TCP query at this point.

RUSS MUNDY: Liman, do you know of anywhere – you cautioned us a couple of times, [in the] TCP measurements, where some of these challenges have been written down or done a presentation or a paper or something that we could go look at and then make reference to for the work party so people could see?

LARS-JOHAN LIMAN: Sorry, no, I don't, but I'll try to contribute some thoughts that come from the experience I had with the applicant guidebook, which, agreed, is a different problem, but it's still made me notice a couple of properties that make things difficult.

DUANE WESSELS: Alright, Russ, I feel like I'm running out of things to talk about with respect to latency. I don't know if we want to move ahead to the next one.

RUSS MUNDY: Well, I think there's one more question in our TBD list that's in the document. We've already agreed to add a couple of things that are primarily dealing with how to describe the numbers of the data that's collected. Are there any other items that should be added to the TBD list that the work party needs to take a look at and try to solve? I

certainly can't guarantee we'll come up with a solution for all of them, but we want to try in this extended work party to collect up what other pieces and individual items we need to try to tackle. So right now, I think it's a fairly good list here and these are specific for latency. But I wanted to open the floor for anybody that had other thoughts that they have on this.

No? Okay. Alright, so we have our things to be filled in for latency then.

BRAD VERD: Paul has his hand raised.

RUSS MUNDY: I'm sorry. Yeah, go ahead, Paul.

PAUL HOFFMAN: So I guess I have the advantage of not seeing all of your faces sitting here, so as I'm staring at my screen, I realized we have mixed up – or so far this has mixed up two things: measurements and reporting. I think that it is reasonable for us to be as clear as we can on everything we want to measure, but we don't have to say now what will be reported, or by whom, because as we discussed earlier, these measurements will be useful to different people, maybe the 037 038 determining group, maybe researchers. So as long as we're saying what is being measured, the, "Oh, let's then do daily medians, no, let's do quartiles, let's do this," is not needed at this point. And in fact, different groups will certainly say, "No, I know how to interpret that data better than that group. So focus [this thing] on what needs to be measured so that all of those

groups have enough in their hands to do something useful I think is a better emphasis. Thanks.

PAUL VIXIE: I agree. There will always be a bigger data scientist than you have, and they should be able to compete on discovering meaning in your runes. So we should really make sure that we do a particularly great job of indicating what to collect, how to represent it, and maybe even a little about how it's stored, and then separately say these are the analyses and potential conclusions we would like to draw from that data," but it should be a separate issue. It's something that OARC really proved with KC's day in the life project, because every year, a bunch of people who were nine years old when it started get into a PhD program somewhere and go, "Wow, look at all this data." And we should really preserve that opportunity.

RUSS MUNDY: Daniel, go ahead.

DANIEL MIGAULT: Yeah. So one of the things that when you measure latency, I think it would be good to also be able to compare that latency regarding other latency, because depending on the network, you may have some differences and it doesn't mean the problem is the root server system. So my comment is coming just after when we say we don't really know what we measure, we have to keep it very simple. I agree with that, but in some cases, especially for the latency, I think it might be good to have

non-DNS measurements to be able to evaluate what we're really measuring.

DUANE WESSLES: Daniel, I think maybe you're proposing something like ICMP measurements to well-known targets that can be used as comparison. Is that sort of where you're going?

DANIEL MIGAULT: Yeah, it's sort of this kind of thing. That's the idea.

DUANE WESSELS: So how do you choose what those targets are?

DANIEL MIGAULT: This is why I just mentioned it's a direction. But yeah, so maybe one of the ideas could be to compare with the top 20 of the most public websites or these kinds of things.

DUANE WESSELS: Okay. Thank you. I wanted to go back to what Paul and Paul just said, which was that it's important to separate the description of how to do a measurement from how those results are I guess maybe aggregated and presented, right? That's sort of what you're saying? Are both of those, do you think they belong in this work party document, or is the work party supposed to only focus on the metric part and leave the second

part for someone else to come in and think about? And Wes has his hand up already.

WES HARDAKER: I agree 100% with what Vixie Paul was saying earlier, that storage is so cheap now, why wouldn't we just record all data for every measurement? The purpose of doing the aggregation is so that we can agree upon a metric that the PMMF or anything else, sort of a standardized metric for how we make decisions based on that data, and that doesn't mean that some researcher can't come up with much more extravagant conclusions based on looking at the raw data and aggregating it a completely different way. But with respect to the PMMF in particular because that's sort of the focus point, they need to be able to say this particular service has an issue here because you've agreed that this is our decision metric, this is what we're going to make our yay or nays against and things like that. And it may change over time, but for now, we could at least standardize of daily makes sense because of [inaudible] patterns and data and things like that. There's some advantages to doing different aggregations. Maybe we want to do more than one. Maybe we want to do aggregations at hourly, daily, weekly, because that's how we as citizens of the world shift over time.

DUANE WESSELS: Alright. Thanks. So it sounds like you're advocating for this work party to put some stakes in the ground and this is what the aggregation should be [inaudible].

WES HARDAKER: Yeah, I think I t's important to – it could be done as an after project, but we're adding more water to the ocean to boil. But I don't think it's that much more water.

LARS-JOHAN LIMAN: The long-term vision I have for this is – and I'm looking gratefully at Naela here – something along the lines of the reports that the CSC received from the PTI. CSC gets monthly reports from the PTI where it reports its performance, and there's a huge number of indicators and numbers in there, and they're all quite well-collected, aggregated, and indicated whether they meet the criteria or not, and if they don't meet the criteria, there is a short explanation on what happened and why and so on, and that's a very rare occasion. It's also filled up with graphs and things.

So this to me is a stellar presentation of this type of data that we're looking at. So if you want something to look at to get ideas from, the PTI report is actually very good. And I would suggest that you do that just to get the mental image of one way that we could chose as our direction.

DUANE WESSELS: Thank you.

RUSS MUNDY: Yeah, and looking at the scope and the SOW, it looks to me like part of the already defined work for the work party is to come up with some set of parameters to use for these, how you describe if an RSO is meeting a minimum level of performance, if the RSS is meeting a level of

performance. So in a way, it looks to me like the work party is by inference trying to describe what the minimum levels of performance of each of these actually are in terms of some measurable numbers. So it looks like it's already defined that we're supposed to be doing that.

DUANE WESSELS: Okay. Thanks, Russ.

WES HARDAKER: One quick note to go back to Daniel's point, because I think it was left behind, and I actually thought it was highly interesting to consider throwing other than DNS measurements into the system. I think I'd like us to consider that at some point. My gut feeling is because I think [inaudible] incredibly important to figure out which aspect of the services you're going to make a decision, is it the DNS service that's the problem, or is it some path in the middle? Traceroutes would be great to record for example. That being said, I'm thinking that this work party should conclude with the DNS measurements first and then maybe do another party on other metrics.

DUANE WESSELS: Okay. Thanks.

RUSS MUNDY: I'd like to ask a question of Wes and Daniel here, and that it is, at least in my experience, significantly more complex to separate out what are the – especially when we're talking latency here in this case – the

network latency from the DNS latency. And as a first step in the process, are we trying to make it too complex by applying this separation rather than just coming up with sort of the basic raw beginning latency measurement and not yet trying to figure out how to separate?

WES HARDAKER: That's sort of why I was suggesting doing it in the future, not now. There is a lot of preexisting work for how to do [inaudible] of latency [versus] over ICMP which is a much earlier portion, that's a kernel pain as opposed to application pain. So there's benefit in doing that, and a lot of people have studied it. And you're right, it makes the data more complex to go look at. That doesn't make it invaluable. But I don't think we have to boil that ocean [yet.]

DUANE WESSELS: Alright. We're a little bit past our allocated time. Oh, Andrew.

ANDREW MCCONAHCIE: Daniel, go ahead.

DANIEL MIGAULT: Okay, hi. So regarding the complexity, it might be more complex to have HTTP request just to compare the measurements we have with the DNS. On the other hand, I don't really see how to evaluate the numbers we have regarding DNS. So we can probably do without HTTP, but it would also maybe require to know exactly which probe and where is its

location. So that may be part of the data that should be provided to evaluate. Maybe that's likely to be networking issue or not.

The case I have in mind is if you have a probe in some islands in the Pacific Ocean, you might expect some higher latency. And it's a good information to know this probe was located in that outlet and not in DC for example.

So that's why maybe the ratio would provide some – but it's more related to the interpretation, but the data we collect maybe should also consider the location of the probe. That might be one way to do that.

BRAD VERD: I like the idea that Russ had of creating a base, a starting point. I also like the idea of what Wes and Daniel are saying with some of these added on features. I think maybe if the work party can just put these in a document and say this is the base that we're working on, and maybe in future, these measurements came up and we address some in the future or whatnot, I think maybe that's the efficient approach just so we document that these came up.

DUANE WESSELS: Alright. Thank you.

DANIEL MIGAULT: I like the idea.

DUANE WESSELS: Alright. Hiro?

HIRO HOTTA: As to the TBD items listed, I think yesterday [inaudible] said that [geographic bias] will be considered by the work party. So it should be written here as a TBD item.

DUANE WESSELS: So geographic diversity of the probes is not described in this section. I think earlier, there was a section that talks about assumptions about a measurement platform and it says something like – it is assumed that probes are distributed, but doesn't really say how they're distributed.

So the work party can do work on that, they can make recommendations that probes should be geographically distributed. We have to figure out how to describe, how to do that, or topologically distribute it also.

HIRO HOTTA: Yes, because why I said this [inaudible] some of the root zone operators system are or maybe will be suited for serving to some geographic region. So should it be the result of the latency [inaudible]? That's why I said that.

BRAD VERD: I'm confused, I'm sorry.

DUANE WESSELS: I think Hiro's saying that some operators may be operated in specific regions, and so that should be taken into account when you're doing the latency measurements.

BRAD VERD: Okay. Again, we're not trying to define what we have today, we have to define what good looks like. So in RSSAC 24, document that we published, we voted on, is out there, we say for a future operator that geographic diversity is important. But are we saying it's not important for the existing operators?

WES HARDAKER: There's a difference between important and mandatory.

BRAD VERD: Sure. It's a discussion point.

WES HARDAKER: There's also a difference between saying the geographical diversity of the entire system is important and mandatory versus –

BRAD VERD: Okay. Please don't focus on important. Let me use the word from RSSAC 24, it's expected.

WES HARDAKER: Of the system?

BRAD VERD: The candidate operator is expected to provide root service from multiple geographical locations. Expected, not important. That's what RSSAC 24 says.

WES HARDAKER: So two in California would meet that qualification.

BRAD VERD: I don't know.

WES HARDAKER: Okay.

BRAD VERD: I'm not trying to argue the system today. So this is what I feel like we get hung up on. We're trying to build a document or a metric system that makes what we have today look okay. What I'm asking is maybe what we have today doesn't look okay. I'm asking that question. I'm not telling you, I'm asking that question. And I feel like I'm getting the answer that what we have today is okay. But we're expecting a new operator to have different requirements put upon them. I find it interesting.

WES HARDAKER: I think – sorry, there was somebody else talking.

BRAD VERD: We're ten minutes over our break time.

WES HARDAKER: Then never mind.

BRAD VERD: Yeah.

DUANE WESSELS: So 15-minute break and we're back at [inaudible]. Alright, thanks.

We're going to try to speed up a little bit. So Russ is going to lead the session. I think we're going to talk about the correctness metric first, which is a little bit out of order in the document.

RUSS MUNDY: Okay, so although we originally thought we'd just walk right down through it, honestly, one of the reasons that I suggested we jump ahead to correctness is in our work party discussions, the correctness metric was actually the closest to come to a relative degree of agreement as to how you would describe correctness, and I don't know how much folks actually had a chance to read through the document before the meeting, but the extremely short version in the RSO correction metric is if the answer to whatever the probe query is comes back and is validated using DNSSEC, then it is considered correct, and whether or not the discussion – we didn't get to the discussion point of what the

specific query would be if there should be some sort of walkthrough, the full root zone by the probe or something different, but determination of what correctness would be was generally agreed that it would be a DNSSEC validation of the response or responses being accurate and measured by DNSSEC.

The other aspect of the correctness assessment was that there would be the current SOA for the current zone, or possibly there would be a need to – when the SOA changes and as it propagates out, that you could be off by a single SOA in the response that pulled the SOA and [inaudible]

UNIDENTIFIED MALE: [inaudible] staleness.

RUSS MUNDY: Did we decide that was part of the staleness? Okay. That's fine. Yeah, you're right. So anyway, it is focused specifically on DNSSEC, and my hope here is that in fact we have a similar view in this work party meeting, but I want to open it up specifically to first address the question, is that an appropriate way to describe the correctness metric?

BRAD VERD: Paul Hoffman [inaudible].

RUSS MUNDY: Paul Hoffman. Okay, go ahead, Paul.

PAUL HOFFMAN: Hi. I sent this to the caucus mailing list a few days ago, but I understand that you're focusing on DNSSEC correctness, but a few people, I think in the prog meeting or in the follow-up meeting, said not everything in the root zone is DNSSEC-signed, the glue records are not, and we should strongly consider measuring those as well because a nonvalidating resolver, some of the customers of the root zone are nonvalidating, in fact it seems like the vast majority of them are, would be affected by for example changing of the glue records as well. So DNSSEC correctness is one way to measure, but exact match is another.

RUSS MUNDY: [inaudible].

DUANE WESSELS: Yeah. Thanks, Paul, for reminding me. I do remember that that came up before, and I apologize for not having it in this document. I think that we can certainly add something around that, maybe as a submetric or something like that. For me, the only thing that's a little bit difficult about that is we have to think about – with DNSSEC it's pretty straightforward, but an exact match or something like that, then you have t o worry about, well, is it going to be the probe's job to know what he expected answer should be, or do you move that somewhere else like a central machine who analyzes all the results and does that matching? But that's manageable. We can certainly tackle that.

RUSS MUNDY: So thanks, Paul. That is a good addition we need to have on our list here, and we should discuss further today, I'd like to, if we can, get some feedback from folks as to first of all, what does this collection of people think with respect to the exact match type of correctness? I'd like to hear from other folks. I don't remember who brought that up there previously. Was it [Robert's story,] or was it someone else? And do we have them in this meeting, who raised that issue?

DUANE WESSELS: Robert's not here, but [inaudible].

RUSS MUNDY: The exact match issue to see if –

DUANE WESSELS: I think it was Paul that brought it up originally. Go ahead, Liman.

LARS-JOHAN LIMAN: I think I disagree with you, Paul. I don't think we should have that in there, because the method that a system provides – and I'm not speaking about the root server system, I'm speaking about the root content and root distribution system as a whole, the method we provide for people to verify the correctness is DNSSEC. And I think we should make a plug for that. I think we should use the method that is provided, we should not use other methods to verify the correctness. And I'm definitely with Duane here that if we're going for exact match, then we need to compare it to something, and that compare data needs

to be obtained somehow by someone. And how do we know that that channel is okay? So DNSSEC is the way forward.

WES HARDAKER: So the root server operator's job is to redistribute the IANA root zone. We can get the IANA root zone from many mechanisms that can be transferred properly. So the question to my mind is not can you do it today. We can come up with a new way of getting the data to verify, but the reality is that the only way to ensure correctness of everything that is being answered is by doing an exact match on glue records. There's no choice. The fact that DNSSEC doesn't cover the glue records is a problem we've long discussed. The reality is we have to have a measurement system that is able to measure the correctness of everything, and the only way to do that is to do a comparison. It's a good point, where do you do that? Does the probe have to do that?

Since the probe's job is already going to be sending all answers back and – don't raise your eyebrows at me, it's true. So I'm stating the issue. If we want to record all measurements real time and we want that to be usable for research and other analysis, then that database's answers, I think that we do want the probe to return all answers back up to the master system so that they can be compared. Because it will be harder to get the probe to download the entire IANA root zone to go check. You're right about that.

KIMBERLY CARLSON: Paul, did you still have a comment?

LARS-JOHAN LIMAN: I'm not sure I agree with you that all responses have to be sent back to some source, because the data we want to report in other cases was just response time. We don't need to send the entire packet back o that thing. Here, we want to report whether something agrees with something else or not, and that's a true/false statement that we can make back. And again, downloading zone file, I know exactly how to do that, and if it's ap robe that I'm not in control over, then we have a problem. But downloading the zone is not difficult. But what it'll do, it'll create another hundreds of thousands of downloading network points. [inaudible].

WES HARDAKER: Just for the record, I did pause in the middle of my sentence realizing I was about to say something that wasn't true. We did conclude that we should send latency and that information back, and we never said the entire packet. So thank you, you're 100% right there.

BRAD VERD: So Liman is correct, the central whatever it's going to be to which the probes report is theoretically and practically capable of just keeping a history of what each zone with each serial number had so that it can look at the answers the probe has observed and see if they match what the zone contained. This would be particularly valuable in the case of nation state interference with DNS operations. If we have any probes in China that end up ever talking to a root nameserver outside of China, being able to bear public witness to the fact that you can't actually get

some information from the IANA root from some places, that is a feature, and I think that Liman's model has merit on that basis.

KIMBERLY CARLSON: Paul Hoffman, go ahead.

PAUL HOFFMAN: By the way, I was not the person who in the work party brought up glue records, I believe. I think it was someone else. But I agree with it, and I agree with it for many reasons, and I think that if we are saying correctness but we are not saying actual matching correctness, part of what Brad had said earlier was we're trying to expose to the world our transparency. If we say DNSSEC correctness, the world can understand that. If we say correctness but we mean DNSSEC correctness, not actual what everyone else in the world thinks is correct, we're going to be confusing people. So RSSAC for this metric needs to decide whether the work party is looking at DNSSEC correctness or what everyone else thinks is the word "correctness," which would be an exact match with a reference.

And there are many ways of doing this. I understand that some of you are trying to design the test already with what we have to pull down the zone or whatever. There are many ways to check correctness or matching that are different. For example, one would be if this same probe is asking 26 other questions, or I guess it's now 52 since they'll be asking over UDP and TCP, if all 56 match, that says I got a consistent view. Doesn't mean I got the right answer, as Paul Vixie just pointed out. There might be a system that is making sure that every DNS query

that looks like this gets this answer, which is wrong, or is not what is in the IANA root zone.

But that is a measurement. Another measurement is to pull down the IANA root zone occasionally. Another is to simply send the query and the response to somebody who says – who we all trust knows what the correct answer is at that time and says it. The answer to what is the A record for this NS record is not a heck of a lot of bytes. Sending that to a central place at a certain time doesn't seem much more onerous than sending, depending on how you like your floating point, two or four bytes, to them.

LARS-JOHAN LIMAN: I'm not going to oppose doing the full check. It just seemed to be unnecessary to me. But you motivated this well, so thank you, Paul.

RUSS MUNDY: Okay. Any more comments on the general topic of correctness? As Duane mentioned earlier, we'll take this forward to future work party meetings and include it in [inaudible] work towards getting in our document.

DUANE WESSELS: I've got one question, I guess, for the group. And maybe it's already on the list, but something Paul just mentioned was – I think he said 52 queries, so is there anyone that thinks that a correctness measurement needs to be performed over all of the transports versus, say, just one of

the transports, maybe chosen at random or hard coded? I'd like input on that.

LARS-JOHAN LIMAN: On a case by case basis, there might be interesting information in there, but I'm not sure for the large-scale measurements, that it's needed, but it could be interesting to see whether UDP traffic is deflected but TCP traffic is not so that you get different answers from different servers. But that's more of a debugging thing than large-scale measurements over time. So I guess that I think it's not strictly necessary. And again, using Brad's words, maybe we should take a step at a time.

RUSS MUNDY: So this is the RSO section, so we'll be looking at each of the RSOs. But am I interpreting what you just said correctly, Liman, that just for a starting point, just doing – for correctness, just using single, perhaps alternating, sometimes UDP, sometimes TCP, something like that would be satisfactory, wouldn't have to do it for everything every time?

Okay. Anyone else have any thoughts or comments on the transport question Duane just brought up?

Okay, good. Thanks. So let's just take a quick step through the correctness questions here for the submetrics. In terms of the interval, we were looking at it being roughly the same starting point as what we were using for the latency interval, I think, and again, none of it's in stone, but one minute for the interval between queries. Anyone have any strong feelings, yay, nay, or otherwise on that?

BRAD VERD: [I think this one could be larger.]

RUSS MUNDY: You think it should be larger than that?

BRAD VERD: [inaudible].

UNIDENTIFIED MALE: Mic, Brad.

BRAD VERD: I don't think this metric needs to be per minute, per five minutes, per ten minutes. Certainly less than a day, but greater than a minute. Somewhere in there [inaudible].

UNIDENTIFIED MALE: [inaudible].

RUSS MUNDY: An hour? Okay. Good. Now, response time on the query before it's considered a timeout. Again, [inaudible] is what we had as the nominal starting point. Anyone? That's a long time for a timeout.

LARS-JOHAN LIMAN: Four seconds is probably fine. The thing that I think is important is that we use the same timeout for all the queries and four seconds feels a good number to me, but just keep it the same for the various tests.

RUSS MUNDY: Okay. Good. We already discussed DNSSEC validation issue as the next one, and the period of aggregation for each RSO, again, one day is the same as what we were looking at for the other type of measurements or metrics. Anyone have any problem or suggestion for a different aggregation period? Yes, go ahead.

KIMBERLY CARLSON: Paul Hoffman, go ahead.

PAUL HOFFMAN: I believe that of the measurements that we're taking now, when they're externally visible to typical Internet users, incorrect answers will set off way more flags than longer latency or things like that. So as we're thinking of aggregation, given that the general expectation is because of what it says in RSSAC 001, that the general expectation is there will be 100% correctness all the time, that when we consider aggregation, that it could be a much shorter thing, just saying, "Oh my god, we found this thing that we truly never expected at this probe at this time."

RUSS MUNDY: Go ahead, Liman.

LARS-JOHAN LIMAN: I'm with you, and I come from a somewhat similar but different angle, which is in light of the recent hijacking problems which have lasted only for a few hours, we want to make sure that we detect things like that. And that should give a signal in the system.

I do agree that monitoring is different than long-term trends, but I agree with you that there are things that can happen which have shorter time spans, and we want them to be visible in the statistics. That said, I don't really know what the appropriate time is, but one day strikes me as a bit long as well. Maybe [on the order of] a few hours.

RUSS MUNDY: So one of the things that has not really reached the discussion point in the work party yet is – and it was suggested earlier that the reporting output be separated somewhat from the identification of the metrics, but I think perhaps reporting is an appropriate thing to talk about here, particularly with respect to the point Paul just raised. Is this type of reporting an error in the context of the zone something that we as a group feel is of sufficient import to flag in some manner, to alert in some manner even though this is really intended to be a system, I think, for measuring the long-term nature of the root servers? What do people think? Brad?

BRAD VERD: How do I say this? I think as we talk about it today, I hope we're not building the monitoring system. I think if 37 were to be implemented,

then maybe they'd build the monitoring system. I think whatever is defined, and I'm guessing most of you would do the same thing here – it's more of a question, would you guys all do what I'm about to say? Which is if I'm going to be held accountable to whatever these metrics are, my guess is that I'm going to go create monitors of my own system to monitor it, probably on a higher frequency than the SLA states.

So I feel like we all as operators would go out and monitor our system based upon these metrics, and then rather than relying on the probe data to tell us something wrong. Because I know where my instances are, the probes don't. The probes know what instance they hit. So I'm going to monitor all of my systems and all of the data for correctness, for latency, for availability, for all these different things.

So I'm a little concerned. Right now, I don't think we should be building a monitoring system. I think if 37 gets implemented, then the PMMF, if contracts for one or builds one, then that's something different.

RUSS MUNDY: Okay. Wes?

WES HARDAKER: I agree completely, Brad. My view is that this document should be creating a specification that could be included into any number of monitoring systems, and they may collect and do other stuff, but we're defining the specification of what the PMMF may – none of this is prescribed – adopt in order to build theirs. And you're right, yeah, the

first thing I'm going to do when this gets published is go make sure that we're monitoring [the woods] too.

RUSS MUNDY: So I think that the question that was brought to the fore by Paul Hoffman is, is the issue, the metric of correctness, one that needs – I'll describe it as a somewhat different level or type of attention than the other metrics that we talked about. Is this something that we want to have be structured somewhat differently? Because it really came out in talking about aggregation period of one day. And if the objective is to have it be 100% correct for each and every day, that would be one way if you could use it as a one-day aggregation period, if the objective is to find out [inaudible] frequently for the output of what this metric requires like once an hour or something that it is correct information, is that something that the work party thinks we should push for? Wes.

WES HARDAKER: Two things. One, we ought to get through talking about all the metrics before we start prioritizing. And two, what we sort of agreed to yesterday was that you're getting into the notion of alerting. So the reality of it is that we're talking about how would a function from 37 look back over three months of data with one-day aggregation intervals and make a decision that something needs to be done. That is very different than I have an immediate problem. So no, I don't think that we need to change correctness to one hour, because that's looking into an immediate problem issue. And it's not that this specification can't be used for that too, but that's not the target.

RUSS MUNDY: Okay, so as of now, it sounds like we have at least a reasonable degree of agreement that using the same aggregation period for correctness is acceptable. I know it's outside of what you said a little bit ago, Paul, but did you want to make any final comment about it?

PAUL HOFFMAN: No, that's fine. Again, I fully agree that we're not using this as an alerting system, or if we are, we have to redesign it significantly. But as long as that is what is explicitly stated, then I think the work party will be able to come up with good numbers to fill in for the TBDs.

RUSS MUNDY: Okay, thanks. And Fred.

FRED BAKER: The comment Wes made a minute ago highlights something that I've been kind of [inaudible] about most of the morning, which is early and late binding of variables. It seems like what we need to say is that there is an aggregation interval, and in a particular use case, the aggregation interval might be a day. And in a different use case, it might be a minute [inaudible].

What we need to say is that it gets aggregated somehow and then we get into discussions of use cases.

UNIDENTIFIED MALE: I think one of the things that I struggle a little bit with in this aspect is if you've got a whole set of measurements, over a day or a week or a month or a year, and you want to characterize correctness, I feel like we need to give advice on how to do that.

So let's say you're measuring on an hourly basis and you've got 24 in a day, if one out of 24 comes back incorrect, does that mean you were incorrect for the day? If one out of [a year] comes back, were you incorrect for the year? Where do you draw those lines, and how do you characterize that? Or do you say I was incorrect?

FRED BAKER: So your use case there is measurement against an SLA?

DUANE WESSELS: Or in a tweet if this data is public, or if the measurement methodology is public, can someone say "X root was incorrect over this time period?"

WES HARDAKER: So when you report failure states for something, if the goal was 100%, you wouldn't report it as just yes and no, you would report it as Z root failed this metric because they were 99.9% correct instead of 100. So you should report the reasons why that Boolean suddenly got set off.

LARS-JOHAN LIMAN: I think Wes said a very good thing there. The thing I want to avoid is that the previous measurements we've talked about aggregating over a day

and using the median. And if you measure every hour, you have 24 measurements, and you can have 11 of them be incorrect, and the median will still say correct, and it will not signal the case that you have an almost 50% incorrectness in your system. So that's why I kind of worry about the median. But I think Wes is hitting the target here. That's where we want to go.

WES HARDAKER: Can I do a quick interrupt? Paul Hoffman, can you tell me how the audio quality is? Others are complaining that it's choppy and sounds like it's being sent through a fan or a voice modulator. Is it the same for you, or is it just a single person?

PAUL HOFFMAN: This has been fine for me the entire morning.

WES HARDAKER: Thank you.

RUSS MUNDY: Okay, I think we are finished with our walkthrough of correctness.

LARS-JOHAN LIMAN: One question. I know we said we weren't prioritizing, right? And I don't want to prioritize.

WES HARDAKER: I don't know if we agreed to it, but anyway, go on.

LARS-JOHAN LIMAN: I'm just curious, and maybe this is a question that we can put on the parking lot for later, which is, what is the order of priority – I guess that's the right word – of data integrity over availability, or over latency? What is the –

WES HARDAKER: Right. No, I get it –

LARS-JOHAN LIMAN: What's the knockdown list?

WES HARDAKER: Can I rephrase your question? Which is, how – eventually, once we define all these, is the PMMF going to make a decision based on the results?

BRAD VERD: Worst case scenario, is it better to give no answer than a wrong answer?

WES HARDAKER: Yeah, so worst case scenario is you fail all the metrics, but I don't know –

LARS-JOHAN LIMAN: [inaudible].

WES HARDAKER: I would argue that we should table that discussion until the end of defining the metrics. But I agree, it has to be talked about.

LARS-JOHAN LIMAN: [Probably fine.] It's in my head, that's all.

RUSS MUNDY: Okay. Thank you. I think we've made it through correctness here in a goodly amount of time, which is great. And the next area we want to go to is availability. So if we could scroll up a little bit in the document, Andrew. Okay, what time is –

UNIDENTIFIED MALE: We have 45 minutes.

RUSS MUNDY: 45 minutes here. So availability is something that we have spent a fair bit of time in the work party discussing, and right now, the availability for each RSO has submetrics of v4, v6, UDP, TCP, so I'm sure Liman will give us the same cautions about the TCP testing. Thank you.

LARS-JOHAN LIMAN: Consider it done.

RUSS MUNDY: And we will then jump into the actual TBD sections, and availability is measured one minute intervals, and that is –

DUANE WESSELS: Russ, I think based on the earlier discussions, [inaudible] I would say – I heard one minute was too often in that discussion, so I think our starting point should be five minutes for this kind of thing.

RUSS MUNDY: I think that's an excellent point from what we heard earlier. I almost feel like an auctioneer, let' start the bidding at 5 minutes. What do folks think about a 5 in availability? Again, we haven't defined the whole duration or the other aggregation piece or whatever, but this would be an individual availability type of query being sent once every five minutes to each of the RSOs by whtever mechanism we're using to do testing. Does five minutes seem about correct in terms of the time frame? Too high, too low? Brad shaking his head.

BRAD VERD: I'm sorry, it just seems way too high. But that's just my gut.

WES HARDAKER: I agree. Availability – and I hate to use the word "is important" because everybody's going to yell at me for hitting on that earlier, but the reality

is availability is something we need to know – the only way to measure availability and get a decent percentage at the end of the day for how available you are is very rapid measurement, unlike latency, which is less likely to change over a frequent period of time.

RUSS MUNDY: Okay, so we now have a one-minute bid, I think. So, is one minute more correct here?

BRAD VERD: I'm going to channel Matt. Matt was here earlier and he said the same thing. He said availability, you want the interval to be less and latency can be higher type of thing.

WES HARDAKER: Yeah, I would also go with one minute. If we have a server that's dropping 20% of its packets, it's going to take us longer to find something like that with a longer interval.

RUSS MUNDY: Okay. I think Suzanne was next.

SUZANNE WOOLF: I'm sort of tempted to back into an answer to this question, which is, what availability do we want to be able to verify from the perspective of the user? And then how much do we have to sample in order to say with confidence that that level from the user perspective is being met or

not? I used to be able to do the math in my head, I can't do the stats for it anymore. But I'm assuming for instance that Duane works with the numbers enough that it should be pretty straight forward to say if you want to have confidence that you will spot an outage or a failure of such and such a duration, here's how often you have to sample.

LARS-JOHAN LIMAN: I just want to make the observation that if you make an availability test, you also get the latency response in the same thing, so I think there is room for a line between the two. It seems stupid to send queries every minute and not use four out of five for measuring the response time.

DUANE WESSELS: It says that here, actually.

WES HARDAKER: Yeah, I was going to say the same thing but in a different way. We're defining a specification that is really saying this is the minimum sampling time, that things may sample more frequently, and yes, you may align multiple measurements together, but this is what we think is the minimum in case maybe one person – you can imagine outsourcing latency to one thing and – why you would do that, I don't know, but the point being is we should be defining the minimum expectations for each of these metrics, and then my guess is the implementation, yes, will roll many of these together.

No, I mean the minimum time period between checks.

RUSS MUNDY: I think Wes brings a good point, certainly as far as how the work party has approached this so far, trying to identify a metric and identify the parameters around that given metric, keeping in mind that when you actually get to the point of really building something, you may do some of these things in a more efficient manner than just each separate stovepipe.

RUSS MUNDY: Anyone else want to comment? So we're sitting at one minute bid on the availability metric, so I think that we'll take that forward as sort of our nominal timeframe for the work party and to work with further.

BRAD VERD: Going once.

RUSS MUNDY: The next parameter is timeout, at four seconds like the others, and there haven't been any objections to that so far, so unless somebody wants to raise an objection, I think we may have an easy one. Timeout is looking good. We've reached what folks seem to agree to there.

Now, the next sub-metric is the period of aggregation. So again, starting point suggested was one day, and in terms of availability, does an initial aggregation point of one day seem to be a reasonable number to start with? I see some nods, yes. Any objections at all? Jeff?

UNIDENTIFIED MALE: The output's a percentage, right?

RUSS MUNDY: Yes, output's a percentage. At least that's, at this point, what's generally the vision for it, I think. Nobody else has asked anything different [inaudible] work party. Duane.

DUANE WESSELS: No, they haven't.

RUSS MUNDY: So it looks like we may have reached [inaudible]. Go ahead.

UNIDENTIFIED MALE: So this is a metric where I think we need to discuss – so the difficulty I have is we've got a lot of data to aggregate, so say if you're doing a measurement every minute, you've got 1400 measurements per day per root server from N probes. So would you envision aggregating all of those into a single metric of availability for that server operator? Or would you want to say something like X percent of probes had Y percent of availability to the root server? So that's another way you could do it.

RUSS MUNDY: So what you're describing instead of just a very simple percentage of availability that includes everything to divide it into one slice would be

availability from some percentage of the probes, and then another slice I guess I'm not sure how that combining that you're suggesting here would actually work if it was a differing amount of availability for a differing amount of probes. So the output then would almost have to be graphical rather than just numeric, right?

UNIDENTIFIED MALE: Maybe, yeah.

LARS-JOHAN LIMAN: Channeling Jeff's comment to my right. Baby steps. Let's go for the combined thing to start with. We can refine this in the future.

UNIDENTIFIED MALE: I'm kind of back to the integration interval. Wouldn't that depend on how often things are going to be checked, the metric is going to be checked and some decision made? And I know if they're going to check on it once a week, then maybe once a day aggregation may not give you the data for the day that you're looking for. But I don't know how that interval that it's checked at gets decided. But that might affect how often you want to [check it.]

DUANE WESSELS: Yeah, I don't know if we're to that [inaudible] discussion. I'm not even sure that this work party would be making those decisions about – that sounds like an SLA check decision, more of a how to collect the data.

TOM MIGLIN: Maybe just leave that fluid, because again, I understand what you're saying and I agree, but if we're going to choose an aggregation right now, that may have to change depending on down the road if you're using a different SLA.

RUSS MUNDY: Yeah. Thanks for this, Tom. And I think that what we're really striving for here is kind of the baseline starting point that we all recognize there's something that going to need to change, because [inaudible] and created 001, 002, thought, oh, these are pretty good. They may not be perfect. And then we found how less perfect they are than we thought. So I think that that may be the way we ought to approach this, is to get a good starting point and go forward. Do we have any more discussion on availability? Because if we don't, we are actually achieved ahead of schedule here. Okay, so why don't we then jump to the next topic since we have –

BRAD VERD: [inaudible].

RUSS MUNDY: Oh.

KIMBERLY CARLSON: Go ahead, Paul.

PAUL HOFFMAN: This is actually discussion of availability and latency. There was just a few interspersed comments there about the measurement system, and I think that that availability and latency being measured together or separately is going to actually be a big thing for the work party.

It might help the work party if the folks in RSSAC who've just been thinking about this came to some agreement about whether those measurements are coalesced or not. We can always run more measurements, but some people will – let's say that we measure availability and latency separately and the numbers don't really line up. Then people will say, "Oh, your measurement system sucks," which could be true, because this is the DNS.

On the other hand, if we push those two measurements on top of each other, some people will say, oh, you should have been measuring this or this.

I strongly suspect that the work party's going to get [gronked] on this one, so a little bit more discussion of, are you folks okay with combining them, or do you feel like they need to be separated, would be really good, especially since we're ahead of time right now.

RUSS MUNDY: Brad.

BRAD VERD: I'm kind of with Liman that if you're doing a DNS availability measurement and you get the latency back, you've got them together. However, there are instances where measuring them separately would certainly come into play if you're doing a latency measurement over ICMP or something.

So I'm not sure what the answer is. I guess I'm just fascinated that the work party would get wrapped around the axel on that, I guess. But maybe you answer it both ways. You say if you're doing the availability check and you get the latency answer, then [inaudible] together, and in the parking lot is latency over other protocols or other means, and that could be an availability – could be separate.

WES HARDAKER: Good point, Paul. And you just brought me to the conclusion in my head why – if I was going to stand up in implementation, I might separate them, because where I record latency from very much affects latency. Where I record availability from, I may want to record from more places. I may want to record availability from a satellite, from Antarctica. I probably don't necessarily think that that's an important geographical location in order to include in my latency averaging, because that's just flat out unfair. But I still want availability to be there.

So there may be instances where from some locations you actually don't want to record latency.

RUSS MUNDY: Paul, could you give a little more description in terms of what you see as the challenge of doing them separately versus combined?

PAUL HOFFMAN: I actually don't think that there's much of a challenge. Really, the question is, what do you folks want? And I think Wes just hit a very good point, which is there could be reasons why you want to have one of those tests done in many more places than the other, because the information you're getting back is for very different reasons. So unless anyone disagrees with what Wes just said, I think a reasonable summary would be measure them separately even though they seem like they're highly related. They only seem like they're highly related in the simple case.

RUSS MUNDY: Okay, good. Yes, Wes.

WES HARDAKER: So that brings back the question of at some point, should we go back to latency and define more explicitly what that means in terms of where you measure from, what is in scope and what is out scope in terms of how you measure latency? I would argue, no, that that's a future discussion item, but it may become important when we start talking about the bar, how do you define what's good if all of our measurement points are from Pluto?

RUSS MUNDY: Yeah, clearly all these will involve the bar. Do we have another hand? Robert Story. Go ahead. Your mute may be on. Perhaps you could type something in the chat room and someone can read it. Would someone read that, ls? I can't see the screen.

Okay, well, Robert, you just jot in what you would like to –

UNIDENTIFIED MALE: [inaudible] latency should be relative.

DUANE WESSELS: If I recall correctly, Robert was suggesting that – this is diving back into the latency discussion, which, do we want to do that? I don't know if we want to do that right now or not.

WES HARDAKER: I'll read it. Are there multiple comments that just happened? It looks like there's Paul right before that too. So yes, alright, I see. And then he added, “Comparing latency to all RSOs versus via just one would give an indication if latency was as local issue or a remote issue.”

RUSS MUNDY: Okay. Good point. So you’ve been watching the time. What is our endpoint now? We have a few minutes yet?

UNIDENTIFIED MALE: [inaudible].

RUSS MUNDY: Okay, so why don't we go on to the next one? We'll keep our latency in the parking lot for reopening some of those questions, because they were especially pertinent to the testing structure that might be required and geographic distribution and so forth. Okay. Staleness. Do you want to do staleness or at least start us on it?

DUANE WESSELS: So this is sort of a proposed metric for measuring staleness, and it's derived from the statement of work that says – here it says RSS, but the idea is that the operators should be serving timely responses. So this is interpreting timely as staleness. The general idea is just to query the root server addresses from the SOA record, look at the serial, and see if everyone's on the same serial number. This can be a little bit complicated because you can envision that a zone update happens right at the same time that you've started doing your measurements, so in order to account for that, there's a simple proposal here which is that you query all the servers for their serial values and you take the largest one to be the current serial number. If not all of them have the same serial number, then you might wait some amount of time, say 10 minutes, and run all those again, and at that point, any servers that do not return the current serial are taken to be stale at that time. This can be a little bit complicated because this requires some centralization infrastructure to take all the results back to and do that comparison, and maybe ten minutes isn't quite enough time to do that, I don't know. Maybe there's other ways to make this happen given some centralization infrastructure.

DAVE LAWRENCE: I also know of at least one implementation which would confound it a little bit because it serves highly dynamic zones, so the current SOA is always the current second.

UNIDENTIFIED MALE: That was Dave Lawrence speaking.

DUANE WESSELS: Yeah. So that's affair point, David. But I guess so far we're restricting this to the root zone, which at this point does not have that kind of [inaudible]. But I think it's important to keep in mind that the root zone does have a certain type of serial number, I guess. It changes, it's got a year-month date and then sort of an index number, and that may affect the way that we measure this. If the root zone serial number was different, then we could [inaudible] different types of staleness measurements that we could do on it, but we are stuck with what we have right now. Yeah, Wes, go ahead.

WES HARDAKER: I have to admit I guess publicly that I never remember what all the parameters in the SOA field are off the top of my head. But there is one for zone refresh on how frequently secondaries should query their upstream for refreshness, and I would argue that 10 minutes [is incorrect] and we should be using whatever value is in that at that particular time for the SOA in terms of determining whether something

is stale, because relying guaranteedly on notifications is always questionable.

DUANE WESSELS: Yeah, I believe one hour is that retry.

WES HARDAKER: I believe that is the current value, yes.

DUANE WESSELS: And I think that's an important thing to discuss, is how stale is stale? Because ten minutes is a pretty aggressive – maybe that value is the right range, I don't know. is it 15 minutes?

UNIDENTIFIED MALE: [Yeah, for a refresh, 15 minutes.]

UNIDENTIFIED MALE: I'm sorry, half-hour refresh, 15-minute retry.

DUANE WESSELS: Okay. So I'm certainly willing to go with that proposal that we use those values as our measure of staleness. However, I would say that the root zone and the system is sort of designed to even tolerate staleness of longer periods than that. The changes are very incremental and change pretty slowly, so even being off by a day, I would say, is not a super

significant problem. So maybe we want to factor that into this metric as well. Paul?

PAUL HOFFFMANN: There have been exceptions to that rule where emergency redelegation or renumbering of some nameserver in the root zone had to be done that day, and a one-day delay was problematic for people. This doesn't happen often enough that we should optimize for it, but I don't want to fully ignore it either.

DUANE WESSELS: Alright. Thank you. Yeah, that's a good point.

RUSS MUNDY: Well, the other thing – and these metrics absolutely are not looking or trying to measure any kinds of activities, but depending on the size and the geographic spots of the probing mechanism that's being used, this particular metric is the one that I think is most likely to show up [inaudible] challenges in Anycast distribution mechanisms that a given RSO uses, because if they try to place it in one of the very slow points and they happen to be the one that is hit by the probes, they will show up more likely to be stale. So I think that's something that folks need to keep in mind, is the propagation of updates to all the instances can have an impact on the measurement associated with this metric.

WES HARDAKER: So it's also probably worth noting that this is perhaps the only metric that also lives in RSSAC 002. This is a self-reported metric from 002, and so this data is already out there. We can envision comparisons between an externally made measurement of this metric and the self-reported one from 002. I guess just something to keep in mind.

RUSS MUNDY: So the content of what's in 002, that's fairly well – at least reasonably well described as to how it is constituted, and I have to admit I hadn't tried to look and do a manual mapping. Is that something that the work party can use as sort of the way that RSOs are supposed to compute that to provide the basis for this particular metric?

DUANE WESSELS: Well, I remember spending a lot of time on it in the RSSAC 002 discussion, specifically around figuring out – basically in 002, you start a timer when you get a notify message, and then you stop the timer when the zone is ready for serving on an instance and then you aggregate all those. And I think we're reporting in the 95th percentile of that or something like that. So it's fairly well-described, but I do remember it being a topic of a lot of discussion. We also in that metric I remember discussing at what sort of precision the value needs to be reported, so I think in the [inaudible] file, it's reported as the number of seconds, but the latest version of RSSAC 002 does make the point that that sort of precision is not really required and that an operator who publishes that metric may – I think it's something like they can round it off to minutes rather than seconds or something like that. So I think it's just to sort of

set the expectation that for the root zone, having the zone propagate within seconds is not a strong requirement and not something that people should come to expect. So on the order of minutes is sort of the normal expectation.

RUSS MUNDY: So this is one that we definitely should try to build on [inaudible] material and use that as at least the starting point for more of the details we're looking for here to fill in.

DUANE WESSELS: Yeah, and we can start to take a look at that. Since it's a very different measurement technique, what we're proposing here is an external measurement. I'm not sure how much [inaudible] we'll find there, but we'll give it a try. Is that remote hand still up, or did it go away? Okay.

RUSS MUNDY: Yeah, but the one concern that I had is that as RSSAC have a document that already sits out there that has staleness identified as something that gets looked at, and if we have another different document that says staleness, we need to be able to provide some kind of reasonable explanation if they aren't basically built the same, why they're different, or we can say they're basically a similar measure taken in a different way.

DUANE WESSELS: Yeah, we can certainly put that in our document. Paul, please go ahead.

PAUL HOFFMAN: I put my hand up and down a couple times because I'm agreeing with Duane and very much disagreeing with Russ. Part of the very basic thing that we have said for the work party is that these need to be externally verifiable, and we cannot externally verify when a root server operator got the notify message.

So it's fine for us to say in this work there is another way that some people are measuring staleness, but not to compare them at all, because really, the externally verifiable, I think, is a very clear line that we should keep. And we talked about this yesterday with the BPQ stuff, which is not externally verifiable but we can say that and we can describe it well, and the same would be true here.

RUSS MUNDY: Thanks, Paul. I think you must have missed the "or" in my last statement. If they aren't the same, we need to be able to explain why they aren't, and that's perfectly fine. There's no problem with that at all.

DUANE WESSELS: Alright, so maybe let's try to tackle some of these particular values here in the last – we've got about ten minutes left, I guess. So for example whereas – well, let's see. So the measurement interval, here I think the proposal was to be measured one hour, so this I guess would be similar to what we're suggesting for correctness, and I think in this example, it would be maybe important to have retries on some of the other

metrics. We would want to have some sort of retry logic on those SOA [queries] rather than just ignoring timeout responses. Wes, go ahead.

WES HARDAKER: We're back to the other problem of at some point, this information will be used in alerting our determination, and at any particular point in time if you measure all of the available addresses and some of them appear stale, you don't know if they're stale until you can measure again. So it may be that we want the sampling frequency to be the same as the [stale] frequency in order to determine that the next one turn out that the next one turned out that one of them is still stale, and it takes more than one measurement in a row to determine staleness. So an hour may be too long. If somebody is stale for two hour measurements in a row, that's actually not able to measure what we actually want to measure, which is 15 minutes or 30 minutes.

KIMBERLY CARLSON: We've got a comment from Robert Story. "If an SOA query is used for some other metric at a high frequency like one minute, we can see when a new SOA is first seen, and then how long after each probe sees the SOA for each letter it queries."

DUANE WESSELS: Yeah, that gets back to the discussion we just had about whether you combiner or don't combine metrics. To me, it makes a lot of sense to do that sort of combining, but I'm not married to it, I guess.

WES HARDAKER: I hope implementations beat the specification and do something better.

DUANE WESSELS: Yeah. I sort of feel like at this point maybe the output of this work party would have almost RFC-like advice like measurements may be combined, not must be combined and that sort of thing. So leave it up to the implementation to do that if they want to.

So I think Wes, you were advocating for instead of where it says one hour here, we would use whatever the SOA retry value is for that minimum level of –

WES HARDAKER: That's my current thinking, yes.

RUSS MUNDY: One question, Wes, about that. Would it make sense to have the normal frequency be one hour? And if one is not getting the latest SOA in the response, then you check again in a shorter time frame?

WES HARDAKER: That's one option, except that then you still won't – what Suzanne said earlier is 1000% correct, when she was talking about math, ironically, and I'm saying 1000%. But if you want to detect an event of size X, you have to sample at a particular rate to ensure that your probability of catching that event is 50% or whatever you want to set your

requirement to, and she even went and found the Wikipedia page about the algorithm, which is –

SUZANNE WOOLF: Sampling theory. [inaudible].

WES HARDAKER: [inaudible] theory, which shows how often you should sample given the fact that you want an 85% chance of detecting an event. So the reality is that even doing it at 15 minutes means that you're not really measuring a 15-minute outage. Your chances of finding that 15-minute outage is if you only perfectly sample along that 15 minutes. So the reality is you should do it every 7.5 minutes or something lower in order to actually determine that.

so what she said was right. if we want to have an X percent confidence in measuring some attribute, we should actually do the math and then come up with the right sampling rate based on the confidence that we want in detecting an event of 15 minutes. This one's a perfect example because we're actually saying we want a 15-minute, essentially, check on the SOAs, whereas the latency one, we haven't defined a number yet. But in this one, we're actually saying, no, we think that would be basically the equivalent of the refresh time. So we can actually go back and calculate this one, and I can probably bet it's going to be a lot lower if we want a 90% confidence rate or something like that.

DUANE WESSELS: I think that's good input. Absent any other input or discussions, I think what I'm hearing from Wes is that detecting an outage equal to the SOA refresh is probably where we would take this. Outage is not the right word, sorry. Staleness. Well, yeah, and then there's always the aggregation and overwatch. I feel like we've discussed that a number of times and maybe have some good advice on that already.

My personal feeling is this is going to be one where we're going to learn from our mistakes. We're going to try this and get some data, and then find out ways that we could do better. So this one's going to be a little bit tricky or exciting or fun.

RUSS MUNDY: So final thoughts on staleness. Or is everybody ready to take a break three minutes early? Leading question. Yes.

UNIDENTIFIED MALE: [inaudible] Individual RSOs expected to [retain] a copy of zole LGRs, label generation rules?

DUANE WESSELS: Were you asking about label generation rules?

UNIDENTIFIED MALE: Yeah.

DUANE WESSELS: No, the root zone operators do not do anything with labels generation rules. That's really for PTI I guess.

RUSS MUNDY: The RSOs serve what PTI hands them. Yeah, Wes.

WES HARDAKER: One quick point back to Robert's earlier suggestion of how to do the combining – or actually lots of discussion around combining. I think the right wording that we want to put in there is possibly important implementation notes that these metrics, this metric may be combined with the previous one.

DUANE WESSELS: All right. I propose we take a break and then after this break we have a one-hour session where we just have one more RSO metrics to go through. So I think we're in good shape and we'll see you back in 15 minutes.

LARS-JOHAN LIMAN: I think we've been through all the RSO metrics.

DUANE WESSELS: Even better. So we can talk about some of the other things that we – maybe some of our parking lot items in the next session. Yeah. All right. Thanks.

Okay so we've got an hour in the next slot and Russ and I discussed and we thought maybe we'd use this time to do some of the what were called the parking lot items. I think maybe the first one would be something that came up on the first session I guess which was I think this issue of what Hiro was talking about the geography may be related to what Robert was talking about what with relative latency metrics. So both of these are around the latency metric. Correct? Yeah. And so I guess just to recap, Hiro was making the point that geography should be considered when doing the latency measurements and Robert had a similar thing, and I think Brad's point was that that may be focusing on the way the system is today rather than we should be looking, designing for the future or thinking about metrics for the future. So I guess I wanted to give either of you a chance or anyone a chance to sort of speak up on this particular topic at this time what may.

BRAD VERD: So I'll just add one other comment that I didn't address yesterday because it just ran out of time, topic kept on going, but I did want to come back to I am – this really to Wes and really I really need to understand clarification because I've seen it in the e-mail and you said it yesterday which is that if there is a – and this is a hypothetical but if there is a requirement or a need for geographic diversity, then that threatens the independence of the roots. I don't understand that. I have a hard time putting those two together.

WES HARDAKER: I apologize. I've had a hard time getting the point across, but I believe that there are things that every operator must do, and geographical diversity may be something – I think it may even be in the list that we agreed to previously that everybody ought to be geographically diverse to some element of. But geographically is not exactly – does I mean the entire earth? Does that mean multiple planets? Does it mean one continent?

So how you how you measure that is tricky, but the point that I've been trying to make repeatedly is we have stated that independence and diversification of implementation of the root server system is an important element that we have today. We could go out and duplicate the best of a design to a whole bunch of addresses and say that that's better. There's no reason why we couldn't do that. But to date we have always said that diversity and independent implementation is a critical feature of the current root server system, and we have to decide what elements have to be the same and what elements are we allowed to vary so that an operator can distinguish themselves from everybody else. And with respect to geographical diversity, I'm sort of on the fence about that and I haven't really come to a conclusion. But the reality is the number of instances I think shouldn't be on that table. That should be every operator should decide how many instances they need in order to provide a valuable service. Whether they need to be geographical diverse so that we have one per continent, one per 500 kilometers, I don't know how to add a measure that and where to draw that line in terms of what's mandatory versus how is that particular operator providing something above and beyond, and then later how do we put that into an SLA with ICANN that may come up? Because ICANN may

want one operator with their particular diverse components and make sure that they stay in that diverse path and a different operator has to be held to a different standard because they're providing something different, they're providing a different aspects to the system. And I know I've said that a lot and Brad, thanks for calling me out on it. I don't know if that clarifies it better. I don't know that I can state it better than that. There was somebody else I talked to yesterday that –

BRAD VERD: Here's my challenge. We all have a number of commonalities that we reach a minimum level at. And so in in my opinion if geographic diversity is one of those, there's some minimum that to me doesn't threaten anything. You can still go. You can implement Windows, you can implement somebody if they can type fast enough responses you can implement it however you want to do it. You just need to meet these minimum set of criteria. So I don't feel that that threatens anything.

WES HARDAKER: Let me ask you a clarifying question. What is geographical diversity buying you when you say that? Because there's multiple elements that it gets you. one of them is just redundancy because if one –

BRAD VERD: So here's my thought process on that and I'll defer to Paul. And that is we are offering a global service to the globe. And I feel that not to our constituents, not to my customers. We offer this to the globe, so

therefore we should present it as a global service. And to me global is geographically diverse. So what that specific criteria is around the SLA? We need to figure that out, I think.

But it's not a service I provide to my constituency and I say it's a global service, because that's not it.

RUSS MUNDY: If you think about the purpose the root server system is to give answers about the root zone to our constituents. Which is potentially any recursive name server in the world anywhere. And as Paul said, there's no standardization for how recursive resolvers decide which server to use and we know that everybody does everything a little bit differently, including completely random. Some of them every time. Right? So if we have an operator that has only instances in a very small geographical location, that operator is gonna get queries, and for some people, the service is going to be considerably different than from other operators, like potentially an order of magnitude difference in latency. So I think this is really the elephant in the room right. I think we need to just speak plainly about this. Wes, with all due respect, if you had more than two instances, would you be making this argument?

WES HARDAKER: So I will be much more diverse at the end of this year. So I am actually not talking about me. So I know it may seem that way because right now I have two. Third one should be online this week, maybe probably next week at this point. That will be farther away. And I have two more planned for that hopefully will be online within the rest of the year. My

goal – I have plans to get to geographic diversity. So this is just an interesting example for where do we want to draw this line. And I think geography – we may want to mandate the geographical diversity as one of those mandatory things where everybody should have one on a continent. What we have in making that statement and being concrete, I'm all for it, and we shouldn't dance around the elephant in the room. And if we want to make that statement, I'm fine. But I don't think that every element – so I was trying to speak more generically, right? This is just the example. I don't think every element that we're defining or everything that we want to measure, everybody should have to necessarily be equal about.

BRAD VERD: And I'm not saying that. I'm sorry, Paul was in line.

PAUL HOFFMAN: There are more than one elephant in this room.

UNIDENTIFIED MALE: [inaudible].

PAUL HOFFMAN: So there was a time when most of the root server instances weren't Anycast and most of them were inside the US, and this was deemed as a big optics problem. I don't care whether it's right or wrong, I'm saying that it was not seen as a big engineering problem, it was seen as a big optics problem, and eventually Anycast got big and popular and

everybody was doing it. And Daniel Karrenberg at some point – I think he may have been the chairman of – I don't know what at the time, made the public comment that more than half of the root nameservers are now outside the US and that shows that we have respect for the fact that the geometry of the Internet is round and has more than four time zones in it. And I think that all of that was true and correct and good and led us in good directions.

We have had a spirited conversation since then about WoodyNet and about the number of you now depending on shared infrastructure to serve your letter from parts of the world that you don't have staff or rack space or equipment of your own. I have always thought that was absurd. Jeff knows that I feel the same way about his deal with CloudFlare. My view is that there is nothing wrong with what B root does today because they will serve very well a certain set of resolves and then the others will make different choices and we will end up with a good diverse solid system that B is a good backup for if people closer than him to some part of the population end up coming to him because the servers close to them go down. That's fine. All of that is fine, and I don't think we should adopt different optics that require all of you to go supplicate yourself to Cloudflare or Bill Woodcock in order to appear to be bigger than you actually are so that we can fulfill a completely artificial requirement for geographic diversity. Do what you can do and do the best you can at it. And that'll be fine.

That having all been said about the optics, let me point out that there is an engineering consideration, and that is that geography means nothing if the path between you and somebody who 50 miles away is to go to a [pairing] point 400 miles away or across an ocean. Then you're not

doing it right. What you need is topological closeness, not geographical closeness to your client population. And if we were going to sort of restudy this, and we had several times yes they started from the position of if we were doing this with a blank slate, if trying to design it today, we could never get to where we are right now. We would have ended up somewhere else. And that is very much true here. We wouldn't have 13 roots. We would have two. The Anycast would be much more like 7706, or dare I say, the [10.internet.arpa] stuff, the AS112 stuff.

With DNSSEC, it really doesn't matter what the source address was of the response. But we're not starting with a blank slate. We're starting with a lot of baggage, a lot of people worried about how they're being seen, how they will be seen how we're going to compare it to new people whether the old [inaudible] could meet the standards that we plan to require of new people, and we're certainly never going to have fewer than 13. That would require major political upheaval that I don't predict. So I think if we're coming down to a question of what are the metrics around locality, which I think is what Russ is trying to talk about instead of all of this, I want to go back to catchments.

I want to say try to be as diverse as you can with respect to both your own previous deployments and other people's previous deployments. If there are already three roots and Singapore then maybe adding a fourth one doesn't help the overall system much and you should aim for something else. Thank you.

MATT: So to be clear, I didn't mention geographic diversity when I made my comment. I came at it from the perspective of latency, which I think is more in line with what you're talking about with [catchments.] I don't care a bit about geographic diversity, and my comment about optics a moment ago was the optics underlie the engineering which was there were too many roots in one geographic location namely within a convenient nuclear bomb blasts of where we sit right now. So yeah. To be clear, where I'm coming from is not any sort of optics. I'm coming at it from an engineering standpoint which is there are 26 addresses and some of them are gonna be chosen randomly by some population of resolvers. And the fact remains we all get queries we're all gonna get queries all the time from all over the world. So I think a reasonable consideration is what is the average latency from all over the world to each of the service addresses. And I think there needs to be a bar and I don't think that bar should be too high.

So for example if someone were within only one location, there are gonna be places where somebody can't get to that one location in under a few hundred milliseconds. So I will go on I will argue that that would not be acceptable, should not be acceptable.

BRAD VERD: Can I follow up? So Matt, I agree with that, but I would like to add that sometimes the reason that it's hundreds of milliseconds has to do with choices that have been made far from you. And if you have let's say 24 equally divided Anycast instances, one per time zone, and you peer openly but there are people in some time zone who don't and who

don't buy transit from anyone who does, you could end up making a huge investment that still doesn't solve the problem for some outliers.

MATT: I agree. That's why I use the extreme example of one because the speed of light is only so fast.

BRAD VERD: Right, so this gets us back to the quartile reasoning that we referenced in the previous segment where sometimes you need to throw out a bunch of outliers before you can actually characterize usefully what you're actually – what your impact really is. That's all.

DUANE WESSELS: This is answering – I guess I've been talking to Wes quite a bit about this, and Matt makes a really good point. One thing is RSSAC is made up of mostly technical people but it has a political role. Otherwise this would just be root ops. And a real political piece of this is when we present something to anybody outside of this room – and this is going to go to either the ICANN board or somebody like that – we were surprised the last time we handed them something. We were shocked that they said how much is it going to cost because we hadn't looked at that. Well, this time, we say here is a proposed how this is all going to work.

First impressions are going to be how many instances do they have and where are they. That's how people outside of this room look at that. So whether it's right or not, it's going to come up. So if we fail to address

and address it well, it's still going to come up so we can't avoid the question. Wes has some interesting points about what are the factors around a root server operator, is there a vitality of the research that makes them stand out or some of them just brute force? Paul and I disagree. I think I made the right decision where we look bigger than we are by what I believe as a correct way, he thinks is a wrong way. I have a huge amount of respect for his opinion, but it's what I chose. That's all the diversity of it. And so if there's an asterisk next to our size, I'm okay with that. And if there needs to be an asterisk next to his size that says terrific research done and root server software developed and all of those other pieces, maybe that's part of it. But we're going to present this and the first thing we're going to say is, “How many divisions has the pope? How many units do you have?”

UNIDENTIFIED MALE: I just want to seek a clarification, because one thing I'm confused about is if like all policy and if you don't care what geographical location, is there a way around reducing latency if you don't truly care about geographical location?

UNIDENTIFIED MALE: What matters most is – I agree with what Paul said earlier – the network path, not necessarily the geographic location. If I have one instance with network provider A in the whole world, anybody who's attached to that network, that network provider typically will send it, route it to that one location. We see this all the time with our instances where we're only connected to maybe a network provider at one location. So while you

might have an Anycast catchment or an instance in the same geographical area, if it's not in the same network path or if there's a prevalent that provider, it's not going to go there. It doesn't magically happen that way.

UNIDENTIFIED MALE: And we see this all the time.

UNIDENTIFIED MALE: So to kind of add upon what Matt was saying as well as Paul, a really good example is G root. We just recently did a design change which kind of lowered our latency due to kind of the network path changes that we've made. So even though those locations remain the same, it just further adds to where topology is kind of what needs to be looked at prior to actual geographic locations.

BRAD VERD: And I come back to number of sites. To me, more than one should be a requirement. I know those of you who went to Prague [inaudible] years ago, there was a presentation that we shared about availability of service and we had done a whole bunch of analytics – we being VeriSign – and we showed the root system, and I think – I'm trying to find the presentation. I can't find it. But the crux of the takeaway was if you guys all remember, we had A root that was Anycasted like to five sites and we had J root that was 100 plus. And then there were the other roots obviously that fell below and above all of that. We're kind of across the whole spectrum. But when it turned out around latency, the number,

like the magic number that we were able to kind of see was there was this very low latency when you hit like between – I want to say it was like seven to ten sights, like eight. It was like this magic number where with one site your latency was way up here and then with just a handful of ones added your latency went down here, and then with 100 sites you were down here. So there wasn't this huge return on going from 8 to 100. But there were other reasons . There's other reasons to do that.

So yeah. I agree topology is ideal. The Internet's not ideal. And unless you've got one site that is connected to every network on the world, I have a hard time with that argument. So I think you need multiple sites. And I think you want to try to connect to as many people as possible. And if the argument is topology then we start talking to carriers. So this goes down this whole different discussion also. So yeah, I think it's more than one but less than 100.

WES HARDAKER: As I said, I think we should come up with what the most important metrics are in order to establish a service and probably determining where the bar is for latency. Matt said a couple hundred milliseconds and that's I think right now I guess the TLDs ought to beholden to 500 milliseconds, and that's quite high. You could do that with one site on the planet. My point though is that it's a slippery slope and that the more we keep saying that everybody must be beholden to these individual stats, we lose independence, period.

And so we right now in multiple protocols are prioritizing latency over all other attributes. That is the most important thing, like web browsers

have been having races for latency, for loading web pages, forever. The whole reason that we're putting caches for – like every major CN in local ISPs is to reduce latency.

And the funny thing about the DNS is that we're talking about the root server system with a two-day cache. And so latency actually matters more for negative answers. But as the deployment of aggressive [insect] gets deployed, latency is going to even matter less for the root. And so it may be there are other elements that somebody would never send up a service today that was a single site but in a nuclear bunker so that when the rest of the operators got wiped off the planet, that one would be the one remaining. But you couldn't do that in multiple sites because it's too expensive. So the instant we start prioritizing and saying your implementation must have multiple sites – which I don't actually disagree with, again, I'm using this as a as an example. But we're wiping out extreme thinking and out-of-the-box thinking because everybody now has to absolutely prescribe to latency being the most important.

So whatever engineering solution they come up with is immediately going to eliminate a whole bunch of extra stuff.

RUSS MUNDY: Paul, go ahead.

PAUL HOFFMAN: Thank you. Although you do need to say Paul H vs. Paul V. But thank you. So this is an interesting discussion, but is it part of the metrics work

party? So if it is, is it sufficient? And again this will clearly – since we're talking about instances, these are not going to be externally verifiable or if they are meant to be externally verifiable, that's a completely new rule that I think is unexpected. But is it sufficient for a root server operator to say the number and geographic and topological location and just have it be self-reported for now?

RUSS MUNDY: In answer to your question, is this actually part of the work party discussion, this is parking lot related discussions that are driven by latency, and the latency has then been extended to discuss a number of geographic instances. And so for now I think we can go on a little bit more because it does affect the end latency things that the work party has to determine.

And I think Fred was next followed by Brad.

FRED BAKER: It seems like there's another aspect of having multiple sites that we're missing. That is that there is a requirement for resilience for the service to actually [be up,] and having multiple locations is a strategy toward resilience.

BRAD VERD: Yeah, I was just trying to echo that. The argument that you were giving was for latency is common, and we lose our independence. I'm not the statistician, but I've been in the meetings with lots of them and I think with one site your availability is no better than 95% or something like

that. And then if you had two sites, you'd get 99.5%, and then with there were three sides you get 99.8%. So if availability is above – going back to the above latency, then I would argue it's not just for latency.

RUSS MUNDY: So I'd like to follow up on Fred's lest comment in terms of resiliency. So far, the work party has not really identified resiliency as a metric. Is that something that you think that the word party should look at?

FRED BAKER: Resiliency might be another word for availability.

RUSS MUNDY: Okay.

FRED BAKER: I'd have to think about that. I think they're related.

RUSS MUNDY: Yeah, I think we have an approach [for] availability but we really haven't talked about resiliency. So yeah, if you would think about it and poking further if needed. So we have another area that Duane and I talked about that we had spent a few minutes on earlier. This is another parking lot item and that is, can we get some input on quantity of the testing mechanisms themselves? And probes is the short name for thjem, whatever the process may happen to be. What are people's thoughts on, should there be some minimum quantity of probes that

are needed that will conduct these set of testings? What is the amount of diversity geographically of the probes or topologically from a network perspective? Should there be perhaps a relationship in the quantity RSOs to the quantity of probes to where you can get enough appearances of things around the world that you can test them? Matt, please.

MATT: A couple of times this morning I've wondered to myself, it seems to me like we're discussing things that surely more than one academic has already considered these topics. And so I had that thought again just now when you were mentioning that I wonder if there's academic literature on this where – if you're this many probes distributed this way –surely someone has researched this topic, I would think. Maybe not in the context of necessarily monitoring name servers but just in terms of overall Internet measurement.

RUSS MUNDY: There certainly are a number of those, Matt, and I'm not sure if you were here at the beginning but Duane showed us a series of slides from some of the work that he did and looking at the data from Atlas probe collections where there's roughly 10,000 of them so that's a big huge number, but one of the things we wanted to get some input on is if there were thoughts from this collective group of people about is it necessary for the work party to try to come up with some guidance or input coming back on where the minimum number of probes, how should they go about being identified? Do we want to use purely stuff

that exists to begin with, whether it's the [inaudible] stuff or the RIPE RCC, something else? You look like you're gonna ...

MATT: Yeah, I guess my suggestion is that would seem to me something – so yes, I think that would make sense for the work party to consider that and come back with a recommendation, but my suggestion is that that would require a literature search that surely people have done a lot of work on this and that that's not original work that the work party needs to do, at least not entirely.

WES HARDAKER: I'd be happy to forward a paper by [inaudible] and some others that actually measure the latencies and showed that you need eight deployed instances in order to get where that curve actually is and it's seven to eleven or whatever somebody mentioned it is.

I don't know. I suspect that you're right about where to put probes has been studied. It's complex though, so I don't know that we finished studying that. Let me give you an example. I had a great conversation at the IETF with somebody from CentralNic where they were explaining to me the geographic issues with Central America where all of the countries in Central America have a pipe back to Miami and not to each other.

So if you put a probe in one country and a server in another, the round-trip time still has to go all the way back to and from Miami. So the problem with measuring latency as the minimum is that we're

optimizing for 90 percent of the population and those others are sort of stuck. So when you're talking about probe measurement, we need one everywhere and every network is probably the only correct 100% solution, and that's impossible.

UNIDENTIFIED MALE: But isn't that also the reality for resolver clients in that country? Their have to go that way as well so that's a valid measurement.

WES HARDAKER: My point is that if we only have one probe in a particular region where there's routing issues – it's also been heavily discussed that India has a completely different routing commercial framework for how they route packets internally. So if you put one probe in India you're not sampling the entire ecosystem there.

So we are just talking about trying to come up with a measurement that will say yay or nay you're good or bad. That is not – if we're gonna come up with something that's the best that we can do, or I should say the implementation is, we're not actually – in this metrics work party we've already decided we're discussing what should be measured and how it should be measured, not where and from. The whole thing I brought up earlier of Antarctica versus the moon in terms of latency is still going to come back and bite us at some point. Implementations of this specification are going to have to deal with that and figure out what portions of the globe are we optimizing our latency measurements from. You can't do it without optimizing for particular portions of the globe over another.

DUANE WESSELS: So one of my ideas that I've been thinking about as we've been discussing is I think it would make a lot of sense if we can define metrics in a way that they can be run on existing platforms, and then we can for example take them to those platforms and start to gain some experience with what things look like. Just to iterate and see if we're on the track and doing things the right way. And I wanted to go back to something that Paul and Kaveh said earlier which was that RIPE Atlas is not the right place to be doing measurements from, which surprised me a little bit because I think it's a pretty good platform. So can you say more about what your concerns are there with doing measurements from that?

KAVEH RANJBAR: So from my point of view, I didn't say it's not a good platform to do it. It's just like I wanted to point out there's a distinction between the eyeballs measurements which are done by probes and we have also data center-based anchors. So you can get two completely different views based on what you want. What Paul mentioned, at least what I understood, was yes, on the eyeball side, on the end user side there are all kind of filters and all kinds of things that are out of our control. So if you're measuring using that, you have to consider those. But that's basically what the end user really sees. If it's full roots, I guess it's mostly what the resolvers see, not what the end user at home sees. So anchor might be better fit for that because they're closer or next to resolvers.

DUANE WESSELS: Okay. Thanks. So it would be totally reasonable to have this set of metrics defined and then take them from either existing anchor measurements or run some new measurements on them and sort of see what it looks like and see if we're on the right path. Did I see a hand go up somewhere? Yeah.

PAUL VIXIE: If what we are trying to measure is something having to do with placement in some space, whether it's topology, topography, whatever, placement relative to goodness, in other words some things are good, some things are better than others even though they might look the same from some perspectives, we're looking for a perspective which accurately differentiates the goodness, then certainly the example of South America will never be equaled in terms of the way to measure the experienced performance of South American economies is to put a probe in Miami. That has always been true, that may always be true. That's true in some parts of Asia as well with regard to [inaudible] and California.

So when each of us has spoken about catchments every time we've mentioned that, the references to hydrology. Where does a raindrop fall? And based on where it falls, can you predict that it will go which way will be downhill for it from that location? And the unusual and uncomfortable fact about this is that our catchments have a non-fixed number of dimensions that can overlap in a way that you cannot see whether they've done that. So that's why yesterday when we were talking – again we were off in the weeds then as we've been today, we

were talking a little bit about how to measure goodness based on placement in some space, some geometry.

The point I made was that it's kind of up to the anycast operator who places a node somewhere to find several independent unrelated to each other unrelated to them places in the local catchment from which measurements can be made in order to correctly observe what a real recursive name server in that catchment would be experiencing in terms of service level from that anycast node. And if that means that we've got let's say of four of us in four different South American countries but we all end up with probes that are hosted in Miami, then that's maybe bizarre, but it might also be correct, and that's why this is very much a subjective measurement.

What your catchment is depends on what your topology is, who you're connected to and where you make those connections. And if we're going to try to standardize this in terms of a metrics workshop that's ideally trying to produce a document that guides people as far as how to measure, we would start that guidance from that subjective point of view rather than trying to determine some objective point of view which does not occur in multi-dimensional geometries. Thank you.

KAVEH RAJNBAR: So just two high level comments not directly related to that, but I think it might be useful for the work party. One is as most of you know, but just maybe to reiterate, Atlas is run by RIPE NCC, they are not for profit and basically neutral. All the data is publicly available. Also the source code, the methodology, all of it is publicly available, and funding is

basically – it has different funding sources so it's not a business, we don't have any income from Atlas, but we secured funding for the foreseeable future so I don't see any issue in the long run. That said, I don't think any of this should be based on one platform, but that's just informational.

Second point, just so you know, in the next 6 to 12 months we are also deploying software probes. Right now we have about – I have to check the latest number but it's always about 8 to 9% of active ASNs which we cover with Atlas right now. Our simulations, we think with software probes we can go easily up to 25 to 30% of IPv4 ASNs, and that should happen in the next six to twelve months.

DUANE WESSELS: So I think we've had a lot of discussion just now and earlier about how hard placement is, especially for these probes, and how I guess it makes latency measurements not objective, to use Paul's phrase. But is there anyone that thinks that this is too hard? Should the work party not be doing this, or we just to do the best we can?

BRAD VERD: It's too hard to engineer for 100%, yes, I agree it's impossible. I think we should be engineering for 80, 85, 90. And I think if you do that it's very possible and not too hard.

RUSS MUNDY: What's that a percentage of?

BRAD VERD: Well, the reference earlier was Internet users or resolvers or – we could define whatever you want, but I think if you have a probe in Miami that covers South America, you're covering a lot of the outliers there. That's all.

RUSS MUNDY: Were you speaking General sense, like we can't have a 100% solutions problem but we could shoot for an 80%?

BRAD VERD: Yeah.

RUSS MUNDY: As opposed to like 80% of ASNs.

BRAD VERD: No. Yeah, your first comment.

KENNETH RENARD: Regardless of the measurement performance, probably not going to be perfect, but still I think very well worth pursuing as long as we can identify some of the problems with the measurements. A mechanism, especially when we're tied to service level whatever, if there is a case where a probe or large set of probes goes down, okay, we want some mechanism to say “Hey, look, we need to declare these outliers. But

we'll be able to investigate more for the purpose of improving our own service than meeting the numbers."

But we realize there's problems in the measurements and we just have to deal with it.

DUANE WESSELS: Steve, can I ask you to scroll down to the very end of this document? There's some material that I copied and pasted from the gTLD guidebook, and I guess I want to be careful and not say – I'm not suggesting this is how we do things, but this is just as a data point of how some other entity is doing sort of similar things. Scroll down just a little bit more to see that third paragraph says that, the placement of probes, probes for measuring DNS parameters shall be placed as near as possible to resolvers on the network with the most users across the different geographic regions. Care shall be taken not to deploy probes behind high propagation delay links such as satellite links.

So that sort of covers some of the things that we've been talking about. The fifth paragraph says the minimum number of probes to consider valid measurement is 20 at any given measurement interval. Otherwise they're discarded. Now I don't know how many probes they actually have. Do you happen to know, Steve? Does anyone know about this system in place? Okay. You're asking.

UNIDENTIFIED MALE: [inaudible].

DUANE WESSELS: Yeah. But I just wanted to highlight this as something that's already out there. Fred, go ahead.

FRED BAKER: If we were going to adopt that language – and that's the question – I think I would want to make one edit, and that is in the third paragraph to not unnecessarily deploy probes. There may be somebody that you can only reach by satellite.

UNIDENTIFIED MALE: [inaudible].

FRED BAKER: There's something, but I don't think it's fair to say that they don't exist.

RUSS MUNDY: But as perhaps a counterpoint to that, they may want to measure from their location, and hopefully will have the specifications structured so if it's an Atlas probe measurement and they have an Atlas probe there, they can they can do that.

DUANE WESSELS: I'm not sure where we should take the discussion next. We've got about 15 minutes left now before the lunch break I guess. Wes has a suggestion or topic.

WES HARDAKER: I'm tempted to throw out as a last comment right before lunch. I don't want to start a dialogue, but I came up with an example of how something we might be ruling out. So imagine we have latency measuring framework up. We have probes that are scattered throughout the world. 20 to 100, really doesn't matter. Imagine if my desire was to take my existing system which has a node in LAX and Miami for redundancy, and I want to optimize for a completely different side of the world. I want to take root zone Raspberry Pis and drop them in the Himalayas and Tibet and Congo and eastern China and Antarctica and northern Canada because I'm trying to hit the underserved population of the world, including deepest Peru. I would never do that if I knew that I had to meet some metric of 50 milliseconds as a latency metric. We'd be ruling that out.

So again, not bad, but just be aware that that's the path we're headed, that we actually are forcing people to become more like each other. So I am curious, and I would love the opinion of why did the new gTLD program pick 500 milliseconds, because that seems high to me. We were talking a couple of hundred. So where did that come from? There had to have been conversations about that.

UNIDENTIFIED MALE: I was going to give my pre-ICANN employee answer and realized that I can't do that. So never mind.

UNIDENTIFIED MALE: I was going to bring it up before but Wes gives a good opening to it. When we're checking for availability and things, we serve some really

underserved areas and our availability as such is horrible, and sometimes our response time is horrible because they're just really bad at keeping things working. The unit in Venezuela can be out six months of the year and there's nothing you can really do about it. So I was sitting here thinking if I am being checked on some number of units and how they perform, I'm going to simply deny the existence of some number of them or [dyke] out the power cord or something, because I don't want to be tested on those marginal units, although Venezuela needs a root server, so what am I doing? It's a really good point.

KIMBERLY CARLSON: Paul Hoffman, go ahead, please.

PAUL HOFFMAN: So unlike my boss who's sitting in the room who said that he would not comment from his pre-ICANN days, during the run up to the new gTLD program, I was part of the team that was supposed to look at how to do technical testing, but we were handed this set of metrics already. And when we asked how did they come up and could we modify them and such, basically the answer was we asked some people. We made a wild ass guess, but now it's too late to fix them.

So I would very strongly recommend against using any values here. But I don't think that that's what Duane meant when he put this here. I think he was indicating – and Duane, correct me if I'm wrong – that these are good things to possibly consider including. But the 500 milliseconds, the 5%, and quite frankly, at the time when – because one thing we asked about was that the probes being placed on networks that were near

popular roots, popular resolvers, we asked, how are you going to know that? And crickets. Not [Cricket Lou.] We heard we heard nothing back. So please don't put too much weight on the work that went into these numbers yeah.

UNIDENTIFIED MALE: Yeah. Thanks. You're right. I was not suggesting to use these numbers or even really suggesting to copy a lot of this. I just found this interesting as a data point, as one way to do things. And maybe this is even an example of how not to do certain things.

UNIDENTIFIED MALE: I'm just going to follow up with what Jeff mentioned. At D, we're in a lot of back words, tucked away locations, Africa in particular. I see that as a feature. That's a good thing and I don't want these measurements to be held against me. So I've been thinking about this, and I don't think those probes are going to be selecting those root instance to query. They're going to be selecting whatever anycast is gonna get them. So I'm not really worried about the backwoods nodes being held against me.

SUZANNE WOOLF: I want to sort of support what Paul said about the background here, partly to emphasize a different point which is that another place numbers come from is in the process of negotiation among the stakeholders. As far as various legitimate interests contend over what's an appropriate set of numbers that everybody can live with, because people do have slightly divergent interests as far as what the number

should be and what people should be held to. Particularly in contracts. So whatever we come up with here or whatever the word party comes up with is, I think we can reasonably assume, going to be subject to further discussion with not only proposed contractual counter parties but basically all the stakeholders and people will come up with things that maybe the root server operators in the room don't necessarily want to be held to or otherwise. There's a number of ways it can go, but I think I think it's important to remember and maybe take some of the pressure off that the work party is not coming up with a final cast in stone set of answers here.

PAUL VIXIE: I agree with what Suzanne just observed. I want to go back a little bit to something Duane was saying a moment ago, because the context of our role in the overall DNS resolution process should inform some of the numbers that come from the places Suzanne describes. So 500 milliseconds is absurd. You would call that a failure in any normal operating environment. But it turns out that we don't want to create a system that only very wealthy people can participate in, and if 500 milliseconds happens to be true because you just you just cannot afford to put your gTLD servers in every economy where they have to be we don't want those people to be excluded. It also turns out not to be a disaster.

Forgive this excursion. In my day job, I see about a quarter million cache misses per second. Our sensors are placed right above recursive name servers. That means I see a tremendous amount of the traffic that is being sent to the root name servers. There is a lot of selection bias

there. The people who are cooperatively operating those sensors with my day job tend to be competent. They have in general placed let's say for example their recursive name server not behind a NAT gateway. That just a coincidence that just happens to be if you're willing to run a sensor it's because you've lived long enough to know not to do some things.

What I see as input to C root is very different. And it turns out that Duane and [inaudible] did a study some years ago using F root data that made it look as though 91% at the time of the traffic hitting F root was unnecessary in some way. It was just senselessly repeated queries from people who clearly couldn't hear the answers or malformed or whatever. It just was a lot of crap. I haven't done any kind of a more recent study, although I hope that the DITL people are updating that study somewhere.

But I can tell you that anecdotally 91% wouldn't be a bad estimate. Certainly, it's above 50% of the traffic we see at C root is crap. And I'm not running a comm server anymore, but I'm betting that the Verisign people see an awful lot of questions from people who are clearly not hearing the answer they just got and have – every tried.

What this means is that for the relative minority of recursive name servers who are going to hear, understand and remember what you tell them, 500 milliseconds is not a particularly high penalty for them as far as getting the information they need to then remember the delegation that you sent them and use it until it expires. The 500 milliseconds to the rest of them who aren't hearing your answers or they're sending malformed questions or whatever is completely non operable. That's

not changing anything because they're going to fail no matter how fast you were.

So we really can tolerate a lot of slop because of our role in the DNS resolution process. We are sending delegations. Very few people come to us hoping for an answer. I know that happens in DNSSEC where they need some metadata. But for most of the queries that we're getting is not that, other than after a key roll. Most of the queries we're getting are being answered negatively, and of the ones that are being answered positively, most are being answered as delegations. We should make sure to keep that in mind when we decide what's important to measure and then later when we decide at what threshold does the red light come on. Thank you.

DUANE WESSELS: Break time. Now is a good time to break for lunch. So thanks everyone for the discussion. After lunch we are going to bite off even a little bit more that maybe we can't chew, which is to talk about metrics for the system, the RSS. So everyone look forward to that and we'll see you after lunch.

[END OF TRANSCRIPTION]