ICANN85 Mumbai | CF – RSSAC Work Session (1 of 4)
Saturday, March 7, 2026 – 13:15 to 14:30 IST

DANIEL GLUCK Hello, and welcome to the first RSSAC work session. My name is Daniel Gluck, and I am the Participation Manager for this session. Please note that this session is being recorded and governed by the ICANN Community Code of Conduct, ICANN Expected Standards of Behavior, and the ICANN Community Anti-Harassment Policy. Please observe the following guidelines to participate in the session. I will also post them in chat for your reference. Actually, since this is an RSSAC session, that is just RSSAC members that are participating, right?

UNIDENTIFIED SPEAKER I guess. And our guests are invited.

DANIEL GLUCK Yes, open observers and invited guests and everything else, just feel free to pop it in the chat pod. We do not need to use the Q&A pod for this. So with that, I will post these in the chat and hand it over to Jeff to get us started.

JEFF OSBORN Hi, this is Jeff. Thanks, everybody. I am honored and thrilled to have Paul Hoffman awake at oh dark thirty, and it is remarkable how few people are here. I wonder if it is going to go this way all week. Interesting. So I am forgetting, what else do I do?

DANIEL GLUCK Hand it off.

JEFF OSBORN I do not have to do a juggling act at all? Okay. Paul, off to you.

PAUL HOFFMAN Good. And it is actually not even midnight, and I live in a neighborhood with lots of undergraduates, so I am not the only one up. Greetings, and so this session, the general session, is meant to be on metrics and possible future work related to metrics. The Admin Committee asked me to do a presentation on where we are with the RSSAC047 implementation. Some of this is actually identical to a presentation that I gave a few years ago because nothing has changed. But at the end of the presentation, we will talk a little bit more about what the measurements and metrics mean now that we have looked at them for a few years. Next slide, please.

Just as a quick background, when RSSAC047 was published, one of the recommendations/requests that were made was, "Hey, could ICANN write a preliminary implementation? Not a real implementation, but just one so that we can figure out would a real

implementation be possible, what would happen if we did and such." This is definitely what I have been calling an informal implementation. The code for it is all public. The data that we are getting is all public. So it is fully written and running. It mostly works. There are some problems with it that I have discussed before, but I will certainly discuss again briefly here. It is collecting and running the metrics for all four measurements. That is, it is collecting measurements, running metrics for all four of the measurements that are in RSSAC047.

When I say mostly, it is this last bullet, which the bullet itself says for correctness checking, a response is tested against all the root zones that were first seen in use the 48 hours preceding the query until a correct result is returned. This is doable, but it is not very easy to code. Everything is fine if you can find the correct result in the last root zone. But being able to kick over that is actually difficult. Again, this is not a reference implementation. It is just an initial one. It actually does not do that checking at this point. There is code in there that looks like it might work, but it does not, and it is marked that way. That is the current status. Next slide.

Really what we find is that even if that was fixable and most other things are working, the question that I was asked was, "Well, what has been difficult in deploying this?" Writing the code was not difficult. It is all just a big hunk of Python. Some of you folks have actually reviewed the Python and made suggestions. Thank you. It is not like it is real hard to do. But the deployment part has turned

out to be quite tricky. And again, this is what I had said a few years ago, which is getting enough vantage points.

The two sub-bullets here are from RSSAC047, which basically says they will be distributed approximately evenly among the following geographic regions, and these geographic regions are defined by ICANN, and that they should be in major metropolitan areas with only one vantage point per metropolitan area. At the time we had the workshop for RSSAC047, this seemed like a good idea, basically because what we want is reproducibility. If your vantage points, say we were using Atlas probes from RIPE, those are flaky. They are in weird places. You do not get reproducible results easily. This is why we chose this. Next slide.

What we found is a couple of things. Latin America has very few data centers. Africa has very few data centers. And the data centers in Latin America, basically Mexico City could be considered part of North America. In fact, for those of you in the room who like to follow routing and such, it really looks like it is in San Diego. This has caused us to not have vantage points every place. And even in some cities, in data centers, connectivity is spotty, and we certainly found that with Africa. That is that even though there are some data centers—I am sorry, Africa other than South Africa, Johannesburg and stuff is all fine—but in some of the other data centers in large cities, within the data center, things will go dark for a while. And a while could be a minute, a while could be an hour, a while could be a week.

We see in some of the VMs that we have there that the VM itself becomes unreachable to anybody for a while, although we can SSH into it. And even when everything is working fine, that is working fine for IPv4. IPv6 can actually fail within a data center. And then the last issue is something that we have had both in Africa and South America, is sometimes the VMs just fall over. And by the way, that is not exclusive to the South. We have had VMs just go away even in some of the European data centers. Where they are working, they are working, we are sure that they are working, and then they just go dark. You cannot SSH into them, and you can tell the VM provider to hit it on the side. They hit it on the side, it comes up, and there is nothing in the logs to indicate why it went down. Next slide.

Let us talk a little bit about the measurements and the metrics from RSSAC047 from the perspective of, okay, now that we have a bunch of data, now that we have looked at all of this, what does this mean? And again, just as a quick reminder, there are for each individual RSI, there are four measurements and an associated set of metrics with each one. Availability, meaning can the vantage point get any response from a particular RSI by sending a DNS query? The response latency on that query, the correctness of the answers, and the publication latency. Publication latency is different than response latency. Response latency is how many milliseconds until the vantage point gets a response back. Publication latency is when I ask for the SOA, and I know the SOA has changed because a different RSI has said, "Oh, look, we are now

on a new zone." How long does it take for each RSI from the vantage points to say, "Look, I have got the new zone too." Next slide, please.

We will start with availability. It is easy to measure, but, and this is a problem, if we only have a couple of vantage points, most RSI have many instances. Some of you have hundreds of instances. And what we find is that the availability measurements just naturally only hit a very small number of instances for each RSI. Really we are in fact measuring the availability of particular instances. Now, of course, some of you do change your routing, you pull an instance up or down, we might see a change there. But in the normal steady state, this measurement is measuring for each RSI, probably just one instance, maybe two.

And of course, since the probes are in data centers, the instances that we hit are often the ones in data centers, even if it is not the same data center. Data centers within a region often have much faster interconnectivity than going out to a data center at a university and such. What this means is that when we see failures here, it is usually much more likely to be a routing failure. A few years ago, we actually looked when there were problems with going to L-root. We would see a little blip, and it would say, "Oh, no." And we would look, and in fact, the L-root instance had been up and responding to everybody during the same window, and we had to chalk it up to a routing failure. Is this a good metric? Well, even if most of an RSI's instances are available, the RSI is still available from everywhere. So availability is a weird check. It would

have to be a combination of most of the instances for an RSI being unavailable and the routing being so badly screwed up that we could not get to any of them. Next slide, please.

Now let us talk about response latency, which is again one of the things that people really love to measure. What we find in the real world in our testing is this is basically the same thing as availability. It is easy to measure, but you are only measuring the round trip times between one or a very small number of instances for each RSI, for each probe. And again, because we are hitting on probes that are in data centers, sometimes you get for some RSIs just ridiculously low response latency. It feels like the probe might actually be in the same rack as the instance. It turns out that that is not the case. We did informally look, and in fact, even if you are not in the same data center, you can get one millisecond between the probe and the data centers. So we are not really measuring response latency for the RSI. We are measuring it actually just between this probe and this one instance.

Is this a good metric? As far as I know, and that is why I put a question mark here, there have been no studies about how having longer or shorter round trips in general would impact a resolver. We know we are getting measurements. We see that they are absurdly low for this testbed. But even if this was being tested more legitimately in some way, even though then you would lose reproducibility, would that affect anybody? Because resolvers generally pick an RSI with low latency. So if a particular RSI, if Q-root has a hundred instances throughout the five continents, we

are measuring them, and they actually put a hundred millisecond or two hundred millisecond delay on every response, so we are hitting that delay everywhere. Would that affect a resolver? Because a resolver seeing that would go like, "Hey, I do not care about them anymore. I am going to go send my query somewhere else." So this is the core question about response latency as a metric is, it is measurable, but is it actually useful for real-world resolvers on deciding how important latency is? Next slide, please.

This is the easy one. It is easy to measure correctness in the simple cases. It is a little bit harder in the hard cases. For those of you who have not looked at RSSAC047 in a while, the tests for correctness are much longer than you might expect, and the reason for that is instead of saying, "Let us just do a DNSSEC validation on the request," we actually send out queries for glue records which are not signed and such, and we look to make sure that the responses have the right R code and such.

Correctness is in fact related to publication latency which we will go to in the next slide. As long as an instance has a relatively recent copy of the root zone, it is always going to be sending out correct answers. We are just not concerned about that, except if there is suspicion that an RSO is acting in bad faith. This was one of the reasons during the development of RSSAC047, especially in the workshop, where people were like, "How would we tell if an RSO is acting in bad faith?" Well, we would look; the only way it really could act in bad faith is sending back bad answers. But even then, an RSO that is acting in bad faith could be acting in bad faith only

to certain locations, only to certain IP addresses, however that would go, maybe ranges and such, which we would never be able to measure because we are sitting in fixed places in data centers. It is pretty trivial for you all to tell where our probes are because you wanted to be sure you could measure that. So an RSO that is acting in bad faith would clearly not act in bad faith for the probes.

Just as one side note, this is the only place where RSSAC047 overlaps RSSAC001, that this metric is measurable, but it is the only one from RSSAC001 that is tied to a service expectation. And to be clear, of course it should be tied to a service expectation. The requirement in RSSAC001, which is repeated in a couple of them, that you must send the correct answers, is the basis of the root server system. So this is not saying, "Gee, this is not legitimate." It is just not clear from what we have been doing that measuring correctness is actually going to find anything, particularly when things go wrong. Next slide, please.

And then the last one here is publication latency. Again, very easy to measure. You send out SOA requests, and you look at when the serial number changes. But as a metric, this is actually contentious because some of the RSOs have purposely put instances out in the middle of nowhere on islands and such, that there are reasons for them to be there. And the ones that are way out there, they sometimes have a known large latency, half an hour, an hour for zone changes. What we have found in the past few years is the public latencies, even of a few hours, are likely to be harmless. And

that is of course because the TTLs in the root zone are on the order of a day or two, depending on which records you are looking at.

A publication latency for an individual instance of a few hours or possibly even a day might be harmless. Also, we are almost never seeing those in our measurements because what we have found is that in fact most RSOs are updating their instances fairly quickly. The flip side of this, which happened not due to our testing, but what we discovered a few years ago when a root server operator just was not updating their things at all, or was not updating most of their instances, is that a publication latency of a few days can actually be damaging to the internet because you start hitting things where the RRSIG expiration dates are running over, and therefore things that are trying to validate cannot validate. And again, that was an accident. It got fixed. We did not observe that at all. Next slide, please.

RSSAC047 also then takes all of the measurements and metrics that are for individual RSIs and comes up with some metrics for the whole root server system. So given what I just said about we are not even sure what it means for an individual RSI, these metrics, what does it mean for the availability of the whole RSS or the publication latency of the whole RSS? When we developed RSSAC047, we threw in some numbers. Quite frankly, we pulled them out of the air. And really the question comes down to why do we have them at all? What would failing mean for the whole internet if, in fact, we did not pass some of them? Next slide, please.

I have got just a couple more slides here, and then we can open up for discussion. Really the question is, are these metrics useful? So far no studies have shown, other than for correctness—correctness is super important—or exceptionally long publication latency affect the recursive resolvers in any measurable way. We often think the ones where it would be the most noticeable would be round trip time, but we do not have any studies to indicate that. We also, because of the way RSSAC047 is structured to make the measurements reproducible, so measuring from large city data centers, it hides any of the places where we are not going to get exactly what we expect, which is excellent service from the root server system. And really what this comes down to is wanting a few local measurements like you can make with RSSAC047 to reflect on a whole instance, or quite frankly from the last slide, on the whole RSS. We want that. We would like to be able to measure some of this, but we have not found that this is actually useful. Next slide, and this is the last slide.

In the years since RSSAC047 was published, some people have suggested more metrics. However, none of the ones that I have seen, and I could be wrong here, none of the ones that I have seen in fact can be measured from the outside. So we have this problem where we are not even measuring the bad instances for an RSI from where we are. But the additional metrics that people said, "Oh, you should be measuring such and such," those would all be self-reported. Now, some of those have gone into RSSAC002 data, but that is completely different. That is data that the RSIs are reporting

on themselves. It is voluntary. There are gaps in it, et cetera. So really the last two slides are what RSSAC should consider going forwards, which is are the metrics that we have got in RSSAC047 even useful? And if they are not, or even if some of them are, are there more metrics that we can do? And that is my last slide. We can leave this up. I do not have a generic Q&A slide. So I will hand it back to either Daniel or Andrew to run a Q&A if there are any questions.

ANDREW Thanks, Paul. Do we have any questions from anyone in the room or anyone online? I see Ken and then Liman. Ken, go ahead.

PAUL HOFFMAN Thanks. This is Ken. One of the things you started off with, Paul, is asking whether these measurements are useful. And there are two perspectives. I think your code, and I think that question is pretty clear that this is—is it useful? Is it practical? Your code I think is practical.

KEN RENARD But we have this vantage point problem, which is vantage point downtime, local network issues to the vantage point, the fact that the limited number of vantage points do not cover all the RSI instances. So the question is, can we have a useful implementation of RSSAC047? Kind of a philosophical question here. I think RSSAC047 is asking the right questions. If you could really get those

pure, accurate measurements, it would be very good for assessing the root server system, but—

PAUL HOFFMAN So wait, wait, Ken. What we get is accurate, it is just insufficient.

KEN RENARD Does it accurately reflect the state of the RSI? So, yes.

PAUL HOFFMAN Correct.

KEN RENARD What it is measuring is good, but can it be useful without—what is the total node count on the internet? How—that many vantage points.

PAUL HOFFMAN Exactly. Can it be useful without putting a measuring system right next to each instance or a much larger number of percentage of the two thousand instances? And even then, can you be sure that that vantage point, which is supposedly next to a particular instance, is actually going to be measuring that instance? For example, some of the root server operators will have instances that are so close to each other that you can put a vantage point right next to instance 102, but for whatever weird routing reason, because it is less than

a millisecond difference, it is actually measuring instance number 107. Does that make sense?

KEN RENARD Yeah. Well, I think this is just a more philosophical question of can something useful be developed? Or, sorry, be deployed to get the answers that we are looking for. But thanks, Paul.

PAUL HOFFMAN To get the negative answers. We have plenty of positive answers now. Yes, everything seems to be up and such, but RSSAC047 is supposed to be looking for problems. Can we design something that would actually find the problems? And so far, I believe not. I would love to be wrong because then we could do more, but I do not think it is a philosophical question. I think it is an operational question.

KEN RENARD So Atlas is a good system to compare things to. It helps with some of these issues because it is such a broad coverage—

PAUL HOFFMAN Mm-hmm.

KEN RENARD —but certainly not statistically diverse, things like that, in terms of—

PAUL HOFFMAN It is also not reproducible.

KEN RENARD Right. And it amplifies some of the vantage point problems and helps with some of the others.

PAUL HOFFMAN Right. For example, I have two Atlas probes sitting here in my house, one at the edge router that is going to my fiber, and one that is actually here in my office, which is behind the router. They get different results at different times based on what I am assuming is routing issues, given that they're both on the same LAN. So if one of them started showing a bad thing, such as it is getting a much larger latency forever, is that actionable? And I am sorry, Ken, you said you had two things. There was the question which you had said was philosophical, that I would say is more operational. What was the second one?

KEN RENARD I was saying there are two perspectives on usefulness. The vantage point part as well as the code implementation. I think the code implementation is much easier to answer, and I do think that it is practical. But thanks, Paul.

PAUL HOFFMAN Okay, great. And I do too. Again, the one part that I punted on is simply that I punted on it. I do not believe it is impossible. I do think that you can come up with a sane way to measure against five days' worth of root zones.

ANDREW Liman, do you have a question?

LARS LIMAN Hi, Paul. This is Liman. I am not sure—it is more of a couple of comments than questions. I always get hinged on measuring of delay times of zone updates because there seems to be an intrinsic expectation that the operator of the servers are the ones who decide, which is not the case. It is actually the publisher of the zone who decides the limits for acceptable delay in the refresh timers and the expire timers. So what the root zone says, it is okay to be behind with a full week. And if the signatures do not map that, then that is actually a problem with the publisher rather than with the operator.

Now the expectations are something completely different than what the numbers say. In numerous occasions, there have been discussions about trying to put these ends together and make things be consistent. But that has not quite happened. I believe that a long and better discussion about expected or tolerable delays should be had somewhere, not here. It is a general DNS

problem. It is not particular to the root. I think that that is something that we should have somewhere.

I also keep thinking on the philosophical layer, what are the problems that this is trying to solve, and how can we build a measurement system that finds problems that users do not find? How can the measurement system be better than the users? I would like to have such a system, so I am not averse to it. But I struggle with this because the internet is so wide and so diverse that it is very difficult to create a measurement system that reflects what the users see and what they need. Typically, root servers do not need to be reachable from end users. They need to be reachable from resolvers. And in most cases, people do not use resolvers in their homes. So there is a lot—

PAUL HOFFMAN Wait, wait. Stop there.

LARS LIMAN Yeah.

PAUL HOFFMAN Stop there. That has changed over time. A larger and larger percentage of edge routers are in fact acting as local resolvers.

LARS LIMAN Sorry, what—

PAUL HOFFMAN This may be regional, but it is certainly apparently happening in the United States that the cable modem box that you have now is in fact either acting as a resolver or acting as a forwarder and a resolver in a way that would drive the folks in the IETF crazy.

LARS LIMAN Yeah, I was just thinking because the most common thing I have seen is they act as a resolver, which are forwarding to a single forwarding resolver, which doesn't change the equation.

PAUL HOFFMAN That has been true—it does not change the equation, but that is only apparently partially true for some router boxes now. And it is really hard to measure.

LARS LIMAN Fair enough. Yeah.

PAUL HOFFMAN Which does not make your question any easier to answer.

LARS LIMAN Absolutely not. And if you toss into the equation the growing use of local root, hyperlocal root, that also changes the equation, the balance, and we do not know what's in the future there.

PAUL HOFFMAN Yeah.

LARS LIMAN So these were just a few observations, so thank you.

PAUL HOFFMAN Sure. And I agree with you. It would be good to have these discussions again. I do not believe we can have these discussions in a way where we will get useful answers unless we do some whiteboarding first about what would a useful answer look like. And that one is really hard, especially because when we did RSSAC047, the useful answers were, well, there is going to be a governance system, because that was only a year or two after RSSAC037 had come out. There is going to be a governance system. They are going to want some numbers, so let us have numbers that might be useful to them.

LARS LIMAN Right.

PAUL HOFFMAN That is, here we are in 2026, and that might feel different. But here we are also in 2026, where every root server operator has, for any given value of zillion, a zillion instances. And that is really what this presentation is about: if you have a zillion instances, I am not measuring that you are a zillion. I am measuring a very small

number. If somehow Anycast went away tomorrow and we were back to 13 addresses, I believe that we would have much more interesting values here. But fortunately, we are never going back.

LARS LIMAN Thank you.

PAUL HOFFMAN Thanks. Yep.

ANDREW And I see Russ online. Russ, go ahead.

RUSS MUNDY Yes. Thanks, Andrew. And Paul, thanks for the presentation. It has been a while since I have seen the summary, and I think you and probably a lot of other people here are aware that this has been a very high interest area for myself, and I was very involved with the first edition of this. And I think there was maybe one other aspect of sort of a fundamental motivator that you had not mentioned that I wanted to bring up.

And that is part of the anticipated results was that there could be an implementation developed that would provide some reasonably technical sound basis to respond to the irregular but frequent complaints that users have of, "Oh, I am poorly served by the root system. I need my own root," blah, blah, blah. In view of the work that you have done on this, do you think that there is

anything so far that would provide a reasonable answer for folks in various parts of the world that are going to continue to come up with complaints of that nature?

PAUL HOFFMAN

No, I do not. Now, again, I could be wrong because we actually have not done any studies. Anything that has looked like such a study has come out with a negative answer, but at which point they do not publish the results, because publishing the results also has—and so I and others have said, "Let us wait for the GWG to finish, because once there is a GS, those questions will go to the GS." And something will happen.

Now, I am not convinced that that will happen, but let me do the contrapositive here. In the years since we've done this, we are hearing fewer complaints from people saying, "I am being underserved." And I do not think that's because we're actually doing the measurements. I think that's just because time has moved on, and people who had made that complaint before had gotten factual answers going, "Well, gee, how bad is it for you?" And folks like ISC, Ray Bellis' tool, not widely used, but you can say to somebody, "Oh, you seem to think that it's bad for you. Please run this tool and let us know." And they run the tool even if they do not understand the responses, although they do usually then spend a little bit of time and they go like, "Oh, this is 50 milliseconds. This is 100 milliseconds." And that's even as an end user. That is not their resolver. So I believe your assertion that we are going to continue

to hear these is true. I hope that the numbers go down. But no, I have found nothing other than the fact that we're all getting older to actually indicate why we might hear that less.

RUSS MUNDY: Okay. Thanks. One other general observation is the vantage points: the location of vantage points, how they are to function, and so forth. That continues to be extremely challenging because it is certainly something that we spent a lot of time discussing as we were doing the first version of RSSAC047, and I do not recall if the update there was as much time. But there was a lot of time spent on it initially.

PAUL HOFFMAN: And I do spend time on that every year, by the way. I look for new places once a year because it is not hard for us to do. I have not found any in a couple of years, so.

RUSS MUNDY: Okay. Yeah, part of what was in my head with respect to this comment and further discussion later is perhaps in looking at the architectural layout of things, it might be something to consider to have similar to open source projects where there can be volunteers, and similar to RIPE Atlas, where it would be relatively straightforward for people to say, "Sure, I want to have a vantage point here, and I am willing to provide the VM or whatever it takes to host it." And with an approach such as that, it might be feasible

to significantly increase the number of vantage points. But the question, of course, also has to be asked, can you get more useful data if you take such a route?

PAUL HOFFMAN: If the vantage points are in data centers, or even if they are small data centers, your trade-off is you are not looking at that many instances, but you're getting reliable data. So can we find a different place to put that line of maybe we get less reliable data, but therefore from many more places, and therefore we would be hitting more instances?

RUSS MUNDY: Well, one of the things that was in my mind for an approach like this would be to work with the major DNS software resolver providers to have code that did it as part of their distributions, where it would be relatively straightforward for the operator at a deployed location to check off an option that says, "Oh, yes, I want to contribute vantage point data to this overall project." That would be in some ways even simpler than fielding an Atlas probe. But just one approach or one idea that could, at least in theory, provide a lot more vantage points with very low impact deployment-wise. So just a thought there.

PAUL HOFFMAN: Yep.

ANDREW So I am going to—

PAUL HOFFMAN Thanks.

ANDREW Paul?

PAUL HOFFMAN Right. Yeah, we are hitting the time thing. So yes, thank you.

ANDREW Yeah. Was I just too close? Okay. Yeah, I was just about to say we're running low on time. Liman, you had a quick comment, and then we need to go on to the next presentation. Just go on? Okay. Do you want to say any closing remarks, Paul?

PAUL HOFFMAN No, that was great. Hopefully as we think more about RSSAC047 in the future, this is useful. I know not everyone is in the room, but slides will be available and I am on the RSSAC caucus mailing list, so this could be discussed there. So thanks very much, and I am not part of the next conversation, if I remember correctly. Is that right, Andrew?

ANDREW That is correct. So thanks a lot, Paul. You are welcome to go back to bed. Thanks for staying up late.

PAUL HOFFMAN I will. Thanks very much. No problem at all. Thanks very much. Yep.

ANDREW So that was that, and now over to you, Christian. Do you have slides you want me to share, or are you just going to talk?

CHRISTIAN WHEELER No, I do not have any slides. It is your show. All right. Well, thanks everyone for inviting me. I am Christian Wheeler, part of ICANN staff, the policy team. I am also helping track implementation of the Continuous Improvement Program work across the community, so that is why I am here today, just to kind help raise awareness about that program, help answer some questions, and maybe help discuss any kind of barriers to starting that work.

But just as a preface, as you all probably know already, the Continuous Improvement Program has come out of ATRT3 recommendations, 3.6 specifically, which calls on evolving the organizational reviews which are mandated by ICANN's bylaws, evolving it from the independent examiners into something that is more flexible and led directly by the ICANN community themselves. So this is affecting the SOs, ACs, and the NomCom to conduct reviews of themselves using a centralized framework and

methodology, but still keeping it flexible enough that the SOs and ACs can adapt it to how they work and the work that they are doing.

After the recommendation came out, a cross-community group was formed to develop what that methodology should be, and they came up with a framework which they published in July of last year. It had gone through public comments. RSSAC I believe had members as part of that cross-community working group as well. And now it is up to the community to implement the framework under their own SO and AC specific structures. And so right now Org is helping the SOs and ACs operationalize that framework, so that is why I am going to help here today.

I think that RSSAC is in a unique position because your representative, I believe Naveed, has already drafted a framework document for the RSSAC with criteria and indicators. I am not going to go into what the framework says unless you would like me to, but in a nutshell, it is basically every SO and AC is asked to show how they abide by these five principles, essentially showing that they are effective, efficient, they meet their purpose, they are accountable, and they collaborate to further ICANN's mission in the multi-stakeholder model. And so the goal of the framework is that each SO and AC would develop three to five criteria that lead up to each of those principles, and as well as developing indicators, some kind of specific metrics, whether they are qualitative or quantitative, true or false, but essentially having data to show

whether or not they are meeting those criteria that they decide on that lead up to those central five principles.

The framework has imagined that it would take place over the course of a three-year cycle that then gets repeated. The ICANN Board has confirmed that the organizational reviews should be deferred until one cycle of the Continuous Improvement Program is completed, and that conclusion would be running around July 2028. Because the framework has been published, various groups are in different phases of their work in starting it. I would say everyone is in phase one of assessing and prioritizing how they want to do that framework. Some groups like the ASO have formally deferred that work until they can get a better handle on, for instance, ICP2, which is taking up the majority of their time, so until they have more bandwidth, they have asked to defer it. So everyone is in different stages of implementing that work.

They are also taking different approaches to how they are beginning that work as well. Some groups are leaning more on staff to help develop a straw man of those criteria and indicators, and then they are reviewing that and giving the okay or changing some things. Other groups are taking a much more hands-on approach, forming working groups to develop the criteria and indicators. I believe the ccNSO had done a World Cafe where they polled their groups in a big central session at an ICANN meeting in Dublin to come up with those criteria and indicators.

There is a big variety of how each group is deciding how they want to start beginning that work. And again, RSSAC is in a unique position because you already have a draft which you are welcome to use as a starting point or start from scratch, but I would recommend at least looking at that document and using that as a launching point. I would also recommend to take a look at that CIP work because it is something that even though you have three years to do it, it can sneak up on you. You do not want to wait too long to start it.

I would also recommend that for this first cycle—and this isn't meant to be a gap-finding exercise or work for work's sake—it is really meant for the communities to look at themselves, and based on their own criteria and principles, to see where they may be falling short in certain areas, whether that be participation, attendance, work, timeliness, to see what are ways that we can address it and work that we can do there. It does not have to be something that is completed by the end of the first cycle. It can just be identified, or we can just say we are continually looking at this stuff.

I would also recommend looking at the work items that you may already be doing, the improvement works that you are already doing, because obviously SOs and ACs are always in a state of continuous improvement. They are always looking at things beyond just these periodic reviews. I would recommend leveraging things that RSSAC might already be doing and taking credit for that through the CIP rather than creating new work. I do not think it

needs to be a big headache, for lack of a better term. But I think that it is something that all the SOs, ACs, and NomCom are doing, and so you do not want to wait too long to start looking at how you want to do that. I am happy to answer any questions about the CIP or any other questions you have related to that kind of work. So thank you for having me here.

KEN RENARD

Hi. Ken Renard. So like you said, RSSAC, Naveed and Aram did our own draft criteria and indicators. Can we please have another acronym? Because we need more CNI. Just the terminology there. So when can we go live with that? Is this due at the end of three years? We can turn in our homework early if necessary or what? And then I think you addressed this. How set in stone is that? It is draft. We as RSSAC can evolve the criteria and indicators. My recommendation for RSSAC is to take that draft and get it sent out to everybody, just to keep everybody familiar with it and discussing it. I think there was a proposed small group to look over that. But getting people familiar as early as possible would be a good thing. Thanks.

CHRISTIAN WHEELER

Thanks. It is a good question. The only real deliverable per the CIP program is that each SO and AC deliver a progress report for public comment by the end of that three-year cycle. So by July 2028, a progress report which is based on that framework. The framework itself, how it is imagined, is that you would have these criteria and

indicators that trail up to those general overarching principles. And so you have that in a draft form already. The next step, if you wanted to take that and run with it, add to it, tweak it, however you want to do, the next step would be finding the results of those indicators.

Right now, as an example, one of the sample indicators is the number of RSSAC advisories published. Right now, that's just one metric. So you would fill that out to see what is that number. You would go through all those indicators, find that actual data, look at it, and say, "Okay, where are ones falling short maybe, or where is there room for improvement?" All the ones that look great, great. There is nothing you need to do, nothing you need to improve upon. If there are areas where this seems lower than it should be, then the next step would be taking a look at what kind of activities might be necessary to bump those numbers up and get that to a place where we want to be. That would be done in the next phase as far as the improvements phase.

The last phase would be reporting on that. Did we get those desired results? Looking at those indicators again. And that's what is published in the progress report. Right now, you're halfway through phase one of coming up with those indicators and criteria. I think phase one is probably the most difficult; the one that would take the longest because you're trying to set the scope. The next ones are actually implementing those improvements that you want to do and then reporting on them for the progress report. Again, you do not need to complete everything in that time. Identifying it

is really good. But that is the only deliverable, a progress report by the end of that.

I think that RSSAC is in a good place with having its draft already. The next step is actually crunching those numbers and seeing also what data you have already, which is even possible to acquire. That's another thing; I've noticed in some other groups, they have a lot of great ideas for metrics and data that we should have without really a lot of thought into how you're going to collect that data. Have you been collecting it at all? Is that something that you want to collect? Is that something that is actually useful? Because you do not want to have data just for data's sake. It is obviously a lot of work for you guys. It is a lot of work on staff sometimes. And particularly if you have a lot of indicators, it can be from 45 to 125, depending on how big you want this effort to be in improvement. So I would just bear that in mind as you go into these next steps with the framework. Lars, you had a raised hand.

LARS LIMAN Hi. Lars Liman from NetNode. I would caution against having measurements that go on volume. The goal of this committee is not to produce a high volume of recommendations. I would be much more happy if we produced one recommendation that was stellar than 500 that are crappy. Writing pages is not what I am here for. And that goes for—you have to watch what you measure, and measuring by volumes and numbers is not always the right thing to do. So we need to tread carefully here, I think.

KEN RENARD Just one comment, Lars, and that is what we as RSSAC need to do to refine that draft. If the measurement is one publication, that may be perfect. Next cycle, 50 publications—well, it wasn't enough. I guess we get to be our own judge and jury, and we grade our own papers. So we will start with A+ and then go.

ANDREW I get too close to this microphone. We have talked a bit about this in the Admin Committee about how to kind of attack this. One of the ideas we had was to have staff do a whole heck of a lot of the work, most of the heavy lifting. I think that is how the GNSO Council is handling this, if I understand correctly. But at the same time, we had some interest expressed from both you and Dwayne, and we have these two wonderful people in the caucus in the form of Aram and Naveed, who are really interested in this topic.

So we have this statement of work that I drafted a while ago, which is kicked around a little bit. I think the discussion for the RSSAC is basically how much of this does the RSSAC want to do in the form of a work party, and how much of this do you just want to put on staff, and how much of this is already done in the form of the draft that already exists as part of the CIP program? We do not have to resolve that now, but I think whatever that comes down to in the end should probably be reflected in the statement of work, which is right now full of a bunch of Google suggestions and in that state of flux that a Google document is before it is really final. Which is

fine, but I think that is kind of the decision the RSSAC needs to make now, which is how much of this do you want to do versus how much of this do you want staff to do, knowing that there are a couple of RSSAC caucus members really interested, a couple of RSSAC members who are interested. But I think I am getting the feeling the majority of RSSAC members are not really all that interested in spending time on this. So how do you guys want to attack this? Lars.

LARS LIMAN

Hi. Lars Liman again. I realized I have a second viewpoint, which is that when I first heard about the CIP program, I was concerned because ICANN already has this plethora of reviews, and I have been part of more of them than I care to remember. At one point when I was interviewed by the ATRT2, I reflected that there's a lot of reviewing going on, and that consumes a lot of resources. The person interviewing me said, "Yeah, we heard that a lot." The term death by review was mentioned, and we really have to be careful about that.

My point is that the CIP program must be used in such a way that we can reduce other types of reviews. Because if we are going to add another set of reviews, the death by reviews comes closer. So we should really work hard to find ways to do higher quality reviews so that we can remove some of the work that is going on right now. That goes in a general sense for ICANN rather than just for RSSAC. Thanks.

CHRISTIAN WHEELER: That is it from me. But I will be tracking all your progress. Happy to answer questions anytime. Look forward to working with you guys more, and happy to step in whether it is also to help support staff here. So thank you for having me.

ANDREW: Thanks for stopping by.

KEN RENARD: Thank you.

ANDREW: I mean, not just based on your comment, but based on the general feeling in the room, I think we want staff to do a lot of the heavy lifting and then present something to the RSSAC. That's what I am going to work on the SOW to say a bit more. I probably won't get to that this week. But I think I have some work to do in the SOW to basically say that staff is going to just put all this together and then give it to the RSSAC and then get the RSSAC's opinion of it.

I do not see anyone jumping up and down saying, "No, do not do this. We want to do it." So let us start with that, and then hopefully we can get to something which actually does reduce the amount of time we are trapped in review hell as opposed to getting trapped more in review hell. All right, one more check of the queue. I do not see anyone there. Do not see anything in the chat pod. So I think

	that concludes this session. For the few of us that are here in person, we're here the rest of the day. Our next session is right after this break for what will probably be a short session on our prep for the joint session with the Board. After that, we are still in this room. Is it safe to leave stuff in here, do you think?
JEFF OSBORN	I think that is sure.
ANDREW	I think so. And then after that, we will be in this room again for a review of review session at the end to hear from them. It is just a presentation to us, not really both ways. But with that, we can stop the recording.
KEN RENARD	You know, Robert froze it.
JEFF OSBORN	Thanks, Daniel.
KEN RENARD	Yes.
LARS LIMAN	Recording stopped.

JEFF OSBORN By the way, I noted what that—

[END OF TRANSCRIPTION]