BARCELONA – How It Works: Root Server Operations
Monday, October 22, 2018 – 17:00 to 18:30 CEST
ICANN63 | Barcelona, Spain

ANDREW MCCONACHIE: … Over 1,000 Anycast instances. And again, I'll talk about Anycast right after this. The root server operators conduct operations independently. They are independent organizations, but they definitely collaborate. They collaborate at ICANN meetings. They collaborate through RSSAC. And they spend a lot of time working with the IETF and other I*organizations as well.

The final myth: the root server operators only receive the TLD portion of a query. This used to absolutely be not true; however, there are some changes. So typically a recursive server out there on the Internet today will send the entire domain name up to the root server looking for the name servers because the recursive server can't know what the destination server actually knows before it asks it. So it just asks it for the entire domain name. It says, "Tell me this," and then the root server operator will come back with, "Well, I only know what the TLD part of that is." So most recursive servers will send the entire query.

There are some changes with something called QNAME minimization. I mentioned this very briefly on some of the privacy enhancements that have been made to DNS. So this is actually slowly changing, but for the most part the root servers receive the entire domain name with the query.

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

Okay. Now this going to be a quick explanation of how Anycast works kind of at a conceptual level. Most people are familiar with Unicast, and I imagine [Alain] spent a fair amount of time talking about Unicast traffic. Unicast is much easier to conceptualize. You have one destination and one source. And you have a single piece of hardware, a single computer at that destination, and it answers directly back to the source. One of the downsides of that, particularly with distributed denial of service attacks, is that all the attack traffic also goes to that destination.

So enter Anycast and what we can do is have multiple instances serve the same data to all sources. So that single source can now hit many different instances. And if there are lots of sources, as in maybe the case of a DDoS attack or just a lot of people wanting to perform DNS queries, they're going to be distributed across many destinations. This is determined by intermediate routing policies, which are mostly out of control of the people operating the servers. But in the end, the sources get the data faster and we can do things such as sync or blackhole the denial of service traffic.

Here's Unicast, and you can see we have our little Internet here made up of circles. We have our one source going toward one destination. Pretty simple. Now here's Anycast. Traffic takes the shortest path to the destination. Now we have a few more destinations. So if we had another source say down here, it would probably hit this destination. So in this manner, the traffic can be distributed across many different destinations. Key to this is that all the destinations have the same IP address but due to a trick in BGP routing, we can have many different destinations all listening on the same IP address.

Now here's one of the main benefits of Anycast. In times of DDoS, here we have our user here again going to this destination because he's closest to it. We have a DDoS attacker going to this destination, and the traffic is sunk. Now, of course, if there was another DDoS attacker here, it would also go to this destination. But due to the distribution of the DDoS traffic, many of the destinations can stay up and active in responding to requests from legitimate sources.

A little bit of guidance for your networks if you're a network operator or especially if you run an IXP, an Internet exchange point, some good guidance for you to think about with regards to root servers is increasing your peering connections. Maybe host a root server instance. Ideally, you probably want to have three to four nearby. There are many different definitions of "nearby."

Another thing to think about is RFC7706 which is an IETF RFC on hosting a root server on a local loopback address. That can also help a lot.

Turning on DNSSEC validation in resolvers. This ensures that no one is tampering with the responses that you get from your DNS queries. The DNS responses can get validated if they're DNSSEC signed, and then you know that they're legitimate.

If you're really interested in this topic, I encourage you to participate in and contribute to the work of RSSAC Caucus where the technical advice is really created. My colleague Steve Sheng will talk a bit more about that right now.

STEVE SHENG: Thank you, Andrew. My name is Steve. I work with Andrew to support RSSAC. I'll give a quick overview of RSSAC and its recent activities.

The charter of the RSSAC, the Root Server System Advisory Committee, is to advise the ICANN community and board on matters relating to the operation, administration, security, and integrity of the Internet root server system. This advisory committee was founded in 1998 after the green paper, the white paper led to the founding of ICANN.

I just want to note here it's a very narrow scope, focus on the root server system compared to other advisory committees.

Just to clarify, the RSSAC that produces advice for the community, and primarily the community and the board, but there are other ICANN bodies and other organizations involved in the overall DNS business. The RSSAC is one of those providing advice.

The second very important distinction is the root server operators are represented inside the RSSAC, but RSSAC itself does not involve in operational matters. So I think that's a very important distinction for understanding RSSAC's role.

Within ICANN, there are three supporting organizations and there are four advisory committees, and RSSAC is one of those four advisory committees that advise the board and community.

In terms of the composition of the RSSAC, it's composed of appointed representatives for each of the operators and also an alternate. So each operator can appoint an alternate member to the RSSAC. There are four or five liaisons to other key organizations which I'll show you in the next slide.

There's the RSSAC itself, and there's also a Caucus which is a larger body of subject matter experts. They're confirmed by RSSAC based on their expertise and statements of interest.

The current RSSAC chairs are Brad from Verisign. That's Brad. And also Tripti from University of Maryland. Tripti is ending her term at this meeting, and the RSSAC will have an election for a new co-chair this week.

The liaisons, I think here to understand is RSSAC concerns itself with the serving of the root. But there are other organizations that they need to closely coordinate or liaise with. So for example, the root zone maintainer, the IANA functions operator which that diagram that Andrew showed you just moments earlier, and then the Internet Architecture Board which is responsible for the Internet providing overall guidance for Internet architecture for IETF protocols.

Within the ICANN, there is liaison to the ICANN board to make sure the information is being passed to the board in a sufficient quick manner, with the Security and Stability Advisory Committee (SSAC), and the other committees.

There are currently 100 technical experts in the RSSAC Caucus. Their statements of interest are public. You can go onto the RSSAC website. They will indicate who they work for, what kind of experience they have, and how much time they can contribute to the RSSAC Caucus work. Whenever the RSSAC Caucus issues a document, at the end there's a section called Acknowledgements that lists every RSSAC Caucus member that contributed to the work.

The purpose is really to have a larger body of DNS experts who bring a diverse expertise. There's [through their work and] there's transparency for who does the work and a framework for how to get work done.

The recent publications, there are three of them: the RSSAC039 is the RSSAC Statement Regarding ICANN's Updated KSK Rollover Plan. The "KSK" stands for the key signing key, which the private key signs the zone signing key which in turn signs the root zone. Part of the DNS practice is the key, the KSK, to be rolled every number of years, and ICANN produced an updated plan to roll the key. And the board asked for feedback from various technical advisory committees.

So the quick gist of the response is the RSSAC did not find any reason within its remit to not roll the key. And subsequently, the board decided to roll the key. And the key was rolled on October 11, and so far we haven't experienced any major problems. So that's RSSAC039.

RSSAC040 is Recommendations on Anonymization Processes for Source IP Addresses Submitted for Future Analysis. The background here is DNS operators, in particular the root server operators, sometimes need to collect query data and submit it put in the place for research analysis. So the case in point is A Day in the Life of the In, the [digital] data, which every year the root server operators collect 48-70 hours of traffic and they deposit for analysis.

Within this dataset there's, obviously, the IP address. Some root server operators are not comfortable sharing to identifying the queriers in those DNS [logs]. Some may be by law prohibited from doing that. So there's a need to anonymize the data. So RSSAC Caucus here trying to find the balance between protecting the privacy of the queriers while also the need for analysis and correlation of the data in some instances. So to that end, they proposed four ways or four algorithms where one can anonymize the dataset and then make a set of recommendations to the root server operators.

RSSAC041 is RSSAC Advisory on Organizational Reviews. In this, they really provide feedback based on their own experience how to improve an organizational review. Within the ICANN, we have SOs and ACs and we also have other reviews that need to be done periodically. So I think learning from each review and improve how the review can be improved, that's really important.

There's the RSSAC Information Session where this will go into much more detail than what I alluded to you today.

There are two current work parties in the RSSAC Caucus. The first one is Service Coverage for the Root Server System. Here what the RSSAC wants is to understand and improve the accessibility of the root server system. So they want to define some indicators for service and then also find some tools or methods to measure that to indicate topographically which areas have poor service. Then you can understand where better to put the root server instances. So I think this is what this study is about.

The second one is Studying Modern Resolver Behaviors. So I think this is particularly studying the recursive resolver behavior, how they interact with the authoritative server that Andrew just showed you earlier and also how they interact with the root server system. So the end goal is to propose some recommendations and also maybe updates to the DNS protocol.

So those are the two work items ongoing.

In terms of transparency, from the RSSAC perspective it has come a long way. We established the Caucus. The Caucus mailing list is open, so anyone can observe the proceedings of the Caucus. RSSAC meets monthly and they meet at ICANN meetings. They have workshops. All those they publish minutes and reports of those. Those are accessible from the RSSAC website. There's also a public RSSAC and Caucus Google calendar which the members can subscribe to, to know the upcoming Caucus teleconferences. RSSAC holds public meetings at every ICANN meeting. The RSSAC holds these kinds of tutorials to disseminate the information about the root server system so that overall we have a better understanding of it. Obviously, the liaison relationships and the procedures are all documented on the RSSAC website.

From the root server operator perspective, the Root-Ops meet at IETF. If you go to root-servers.org, they publish the meeting agendas. All the root server operators, RSSAC defines a set of statistics that each operator is to collect and publish. For example, the number of queries, the number of IPv4 and IPv6 queries. Those are published every day, and all the operators are publishing those queries on the root-servers.org website. You can download those per operator. You can compile them together, and you see that for the whole system. They also have individual web pages, and they collaborate on reports for major events. For example, when a major DDoS attack happens, they collaborate on the report. All of those are available on the root-servers.org website.

I think with that, here is more information to find if you have questions about RSSAC, about RSSAC Caucus. But with that, I'll complete my presentation, and I'll invite the root server operators to come and sit in the front and then we'll open for questions and answers. Okay, so any questions? Go ahead. You have the microphone in front of you.

MOHAMED ABUABED: Hello. My name is Mohamed Abuabed. I am from the Fellowship program. Regarding the modern resolver behavior study you mentioned, isn't it out of the scope of the advisory group? Because you're supposed to work only with the root servers. I know the resolvers eventually will communicate with the root servers, but I had multiple discussions during the couple of days ago, and I feel it's out of scope.

Correct me please because I have another question. It depends on the answer.

ANDREW MCCONACHIE: Okay, so Fred and then Wes.

FRED BAKER: Yeah, Fred Baker, ISC and the work party group shepherd for that particular group. No, it's not out of scope because it's about the central operation of the service. The resolvers talk to the root servers, so it's part of our remit. Now our question there is in essence improving the reliability of the DNS as seen by a user, so we want to understand the resolver behavior. Some of the resolver behaviors are less than optimal, and so maybe we can produce a recommendation to resolver operators that would help them to improve the reliability of the service. That's the fundamental thing we're targeting.

ANDREW MCCONACHIE: Thanks. Wes?

WES HARDAKER: So in part it's my fault, right? Because you and I had a conversation earlier, and I was one of the people who said depending on what you're going to study with the resolver, it falls in or out of scope within RSSAC's remit. But the reality is that this project really started from not understanding how resolvers interact with the root system. And, yes, resolvers interact with a lot of systems. They interact with COM and a whole bunch of other stuff. But the reality is that one of the most complex pieces is how they interact with the root system, how they pick the root servers. And that's really where this study is headed, not to study every aspect of how a resolver behaves.

ANDREW MCCONACHIE: Thanks. Do you have a – okay.

MOHAMED ABUABED: Well, I'm not blaming you. Our conversation adding much value for me. I don't know if this question has to be here or in your meetings, but why don't you extend the scope to the resolvers also? As a working group or advisory group, you have to look at the root servers and the resolvers also. Why don't you think about it like this? Because I know it's a huge difference between root servers and the resolvers, but they are interacting with each other too much. So why don't we make it one bigger group or advisory group which handles both of them?

WES HARDAKER: I think the answer there is that we'd like to believe that many of the resolvers operators are in the Caucus or giving the call to the Caucus in order to produce some kind of a result. Going around and saying, "Gee, I've got some random address and WHOIS says it belongs to whoever; let's [knock, knock, knock] find that resolver," that could be a lot of effort in its own right. I'm not sure that we want to go through that level of effort.

So what we are including is the people who maintain the software that is used by the resolvers. We're using resolver operators that want to show up in the Caucus and would like to contribute. [inaudible] you're more than welcome. We're not going to go dig up everybody out there.

ANDREW MCCONACHIE: Thanks. Bret?

BRET FAUSETT: Yeah, just to add to that and to your first question, I think our remit is not to boil the ocean. The whole resolver community is rather large but, and I think it was you day before yesterday that might have asked the question, what server am I talking to? How do I choose that server? Or somebody over here did. Okay. But it's that question that basically drives the conversation around the resolvers because the root server operators have nothing to do with that. That is all done by the resolvers. So understanding that behavior and documenting that behavior so that we can share with our constituency is in our remit and does help the community here, which is why it was on the work list and why it was prioritized.

STEVE SHENG: Thank you. Liman and then then gentleman on the right.

LARS-JOHAN LIMAN: Hello. Lars-Johan Liman from Netnod. I would also like to fill in that the basic relationship between the resolver and an authoritative server of any kind is guided by the protocol, and the protocol is developed by the IETF. So that's another place to have these discussions. But I think it's perfectly within the remit to look at. We operate root servers. We need to understand how the clients come at us in order to provide good service to them. If they implement the protocol in different ways, we need to adapt to these different ways to provide the good service. If we don't understand it, we can't provide that good service.

UNIDENTIFIED MALE: [inaudible] here. I'm representing a small IXP in Switzerland, CHIX. The question actually is, since you're encouraging to run an instance, who do I get in touch with?

STEVE SHENG: Good question. Who do I get in touch with? Okay, start with Wes and then Fred. Go ahead, Wes.

WES HARDAKER: Me. It actually turns out we were just discussing placing an instance in Switzerland or looking at Switzerland as something, so it's fascinating that you're mentioning that. So, yeah, come talk to me afterwards. I'll give you a business card, and we can get in touch.

STEVE SHENG: Fred?

FRED BAKER: Well, if you go to root-servers.org, you can find the websites of each of the operators. I know my operator has a "here's how you contact us." I would imagine they all do. So you can go there and say, "I would like to have an instance. Talk to me." But Wes is a good start.

STEVE SHENG: Bret?

BRET FAUSETT: In addition, the e-mail up there which is ask-rssac@icann.org is RSSAC a year ago or so agreed to act as the window into root operations in the short-term so that if you have a request for that, you could send that to ask-rssac@icann.org and we can forward that on to all the operators who are currently deploying instances around the world and basically connect you to them. And then RSSAC would step out because that is an operational thing. We're just doing this strictly as a convenience until there's some other mechanism set up that's already being talked about.

STEVE SHENG: Thank you for that. Any other questions. I'm sorry. You first and then you.

UNIDENTIFIED MALE: First of all, thank you for that good introduction. My name is [inaudible] from Korea. I am a telecommunications lawyer not a techie. So I am quite curious about the big picture of the system. I have two questions.

Number one is, who really governs the root service? ICANN, IANA, or the U.S. government? That is question number one. My question number two is, I saw some list of root server operators which includes some universities and blah, blah, blah, and there is also ICANN as an operator. Are they paid [off] for this service from registries or registrars, or is it just [free of] service?

STEVE SHENG: So there are two questions. First question, who governs the root server system, and second, do the root server operators get paid for their service? Who would like to answer? Brad, go ahead.

BRAD VERD: Who governs is an interesting question. It really is. Let me answer that in a couple different ways, okay? First, it's not IANA. IANA is a function. It's the IANA functions operator, so it's not even a governing body. The IANA functions is contracted with ICANN right now under PTI. So does ICANN govern the root operators? Today, no. The IANA is responsible for the unique namespace which creates the root zone. And all the current root operators have agreed that they will serve the one root zone that comes from IANA.

There's a document that RSSAC has published recently, RSSAC037, that is a proposal for a governance model for the root servers. This has been a long outstanding issue within this ecosystem, and RSSAC037 is the first attempt to have a discussion around how to create a governance system. First question.

Second question was what again? I'm sorry.

UNIDENTIFIED MALE: Root server operators have been paid for providing this service?

BRAD VERD: Oh, paid? Yes, sorry, I should have remembered that one. No, root operators are not paid. It is all volunteer basis. If you go back and read RSSAC023, the history document, it is a very detailed document that explains how we got to where we are today, and it's all based upon organic growth of the Internet. At the time, the Internet was growing. It was like, "Oh, we need more root servers." I think it was in '81 we had 7. In '93 there were 8. In '98 we went from 8 to – is that right? I got that wrong. Somewhere in there we went from 8 to 11 and 11 to 13. But it's all documented very clearly in the document, and there were reasons for each of those as the Internet grew.

Then at 13 it stopped. Why did it stop? It stopped because at the time there was a packet size limitation for the protocol UDP. Basically, you couldn't send a packet over 512K, so 13 was the max number you could put in. That's why it stopped. Since then technology has overcome that limitation in a number of different ways. We could do more. We don't need to do more because of Anycast. Anycast, you've taken those 13 identities and we now have 1,000 instances of them that are all identical all over the world. So we are able to scale in different ways. Thank you.

STEVE SHENG: Thank you. Wes?

WES HARDAKER: Let me just add one more because you actually pointed out that there are different types of operators. I work for one of the universities. I'm Wes Hardaker from the University of Southern California Information Sciences Institute. That diversity in operations is one of the things that we consider a great strength. It's a money sink. It's hard. I'm one of the smaller, probably, budgets on the system, but we try and contribute in different ways.

We're primarily a research organization, so one of the ways that we try and contribute is not necessarily on a gigantic system but contributing in other ways. For example, we helped a lot recently with analysis for the KSK roll that just happened two weeks ago, and we were actually able to help some other resolvers operators out in the world fix some of their issues. So we contribute in different ways, and we think that diversity is extremely powerful. If you look at 23, it talks about that diversity and adding different people from different regions.

STEVE SHENG: Thank you. I have a gentleman here, and then after that you.

UNIDENTIFIED MALE: Thanks. You've talked about the current work parties on service coverage of the root server system. So if I understand correctly, you're going to talk about optimizing the instances that are in the [inaudible] root service. Is that what is going to be covered?

STEVE SHENG: Not so much.

UNIDENTIFIED MALE: The one thing is also when did this start, this work party?

STEVE SHENG: Liman, go ahead.

LARS-JOHAN LIMAN: I am the shepherd for this working group. First, it has just started like a couple of weeks ago.

UNIDENTIFIED MALE: Oh, okay.

LARS-JOHAN LIMAN: So it's just brand new. The second thing is that we are doing things in several steps here. We are looking for topological areas in the network. I'm not saying geographic here. I'm talking about topology, which means how the operators are connected to each other and how traffic flows between the operators.

What we are trying to identify are areas where the service for the users or for the resolvers is truly bad so that from a technical standpoint it doesn't function well. But in order to do so, we must first define what do we mean by bad. To define that, we need to first define what can we measure. What is technical properties [inaudible] makes things work worse? So we need to define what to measure and then how do we measure this and then try to find levels. Say, "Okay, if it's above this level, it's functioning well and below this level it's not functioning well."

Once we've decided that, we need to create tools that do these things for us. Then we have to deploy the tools to see, "Okay, so now can we possibly identify some areas where it doesn't function well?" So that is a fairly long ladder of steps here to reach the goal of this working group, but we have just begun. If you're interested, please come and join us. We're looking for people.

STEVE SHENG: Yeah, go ahead.

UNIDENTIFIED MALE: Follow up on that. When we design telecom networks, there's a lot of dimensioning activity that happens and also the geographical spread of the equipment. In terms of like I'm from India. So when we talk of mobile networks, for example, we look at where would the network elements be placed geographically and then what will be the dimensioning of that depending on the traffic capacity. So looking at the Internet system, we don't really have an idea about what should be the geographical spread of the DNS instances, how many should be there, what should be the optimum number or something. That's something that we are curious to know about.

Recently, we had some discussion with the ISPs regarding implementation of DNSSEC, and a couple of ISPs had implemented and withdrawn. They reverted back. The reason for that is the quality of service issues that they have faced, because of which they are not keen on implementing DNSSEC.

So one of the things that comes to mind is, is there a relationship between implementation of DNSSEC and the instances of these resolvers that are there? Instances, multiple instances? Because when I [put] DNSSEC, I think there's a latency issue that comes in because of the encryption and all of that. That could be one of the reasons that they are not implementing DNSSEC. So is it that if I increase the number of instances, I decrease the latency so that it compensates for the implementation of DNSSEC so that I come again to the same benchmark of the quality of service? That's a thought that comes to mind. I don't know if you can tell me something on that.

LARS-JOHAN LIMAN: I think the short answer is no because I don't see a relation between the query delay and DNSSEC, as long as we're talking about reasonable changes. It will take less than a second to go around the world, to send a packet the entire way around the world. So as long as we're staying within India or even within the Asian region, the response time from a root server does really not impact DNSSEC.

DNSSEC has a good number of failure modes. There are so many ways you can have problems with DNSSEC, but response time isn't really one of the important ones. If you have problems reaching service at all, if you have a network which just goes up and down [and flaps], that might be a problem. But that's a different type of problem. It's not a DNS problem. It's a network connectivity problem. But the response time is not really connected to DNSSEC, no. [inaudible]

STEVE SHENG: Okay, we'll go with Wes first.

WES HARDAKER: A couple of things. One, I've been heavily involved in DNSSEC for over a decade at this point, and I have talked with browser vendors and all sorts of vendors that really care about speed. So speed is an issue and latency is an issue. And when people really care about latency, they're willing to give up security.

So the reality is that it does take a little bit longer to validate DNSSEC answers as opposed to just asking an answer. However, one of the things Liman said that was 100% correct is that it generally won't have an impact on users – generally. But the very first instance that somebody asks the question, they're going to have to go out and not only are they going to have to go out and get answers for, "What's the address that I'm looking up, but what are the additional details for how I verify it?" So the packets get a little bit larger. It takes a little bit longer to do. There are more roundtrips and things like that.

But remember that DNS is a very caching system. So the next user that asks that same question, none of that has to be done again. So it's very hard to measure the latency because it's different from the first person that asks to the second.

If you read RFC 7706 which talks about how you keep a local copy of the root zone, one of the benefits that outlines is actually not the positive answers but the negative ones. Because if you're asking for some random word that doesn't actually exist in the DNS, sometimes that negative answer lookup – again because that takes multiple lookups on DNSSEC to prove to you that it doesn't exist – that can be a little bit longer, especially if you're beginning to look up multiple things at once.

So 7706 talks about if you're on a bad network, you can deploy a local copy of the root zone. I actually have a project that I run out of ISI called LocalRoot. You can go to localroot.isi.edu and it will help you stand up a local copy of the root zone for your independent network. So you can provide a local copy right there, and that may speed things up. But the reality is there are, I think, seven instances within India right now. So deploying more is actually not going to help you necessarily too much within a geographical region.

And then third and finally, remember that geographical regions have nothing to do with network performance. It's network topology that matters. And there's famous instances of stuff going in and out of the United States and Canada, for example, where to get from one point in Canada to another point in Canada, you have to go through the United States because there's no direct route to get there. So that type of thing exists all over the world. I'm sure it exists in India too. I don't know the Indian boundaries as well. But remember that topology matters more for latency than geographical or political boundaries.

UNIDENTIFIED MALE: But one thing about the instances is it's not really expensive to have an instance. So let's say that your recursive resolver and the instance of the root [inaudible], they're close to each other, would that solve matters?

STEVE SHENG: I think we can take that offline.

UNIDENTIFIED MALE: Yeah, you can take it offline. Yeah.

STEVE SHENG: That also depends on peering arrangements. Brad?

BRAD VERD: Yeah, a couple things. The answer is maybe. There's no black and white answer to your line of questioning here. I will say that comparing deployment of an ISP to deployment of root servers is really apples and oranges. They're not the same thing. If I'm deploying an ISP, I have full control of where I'm sending packets and in which direction and how I'm getting packets because I control all that stuff on the network. Root servers don't. We don't control where the packets come from and how we get it back to you. So, for example, I know because we've had this discussion with a number of Indian people at a bunch of the peering exchanges that the peering in India acts a little differently than in other countries. And the challenge that Wes described is very accurate where places in India you actually leave India to come back into India. And that has nothing to do with root server. We can't help you with that. That is a network topology challenge that is increasing latency to your system that any number of root servers will never solve.

STEVE SHENG: Okay. Sounds good. Any other questions? Okay, going once, twice. With that, thank you for coming. And I thank you, the root server operators, for answering the questions. Thank you.

UNIDENTIFIED FEMALE: Thank you, everyone.

[END OF TRANSCRIPTION]