MARRAKECH – How It Works: Root Server Operation
Sunday, March 06, 2016 – 15:15 to 16:45 WET
ICANN55 | Marrakech, Morocco

UNIDENTIFIED MALE: Interesting. Root ops and RSSAC put this out in Dublin at the last meeting and it was incredibly interesting. So if you have any interest in the root server and root server operations, you're in the right room. And we have Lars-Johan Liman from I-Root server today. I'm sure you have a day job but in this instance – no pun intended – you'll be our Root Server Operator. So I'm just going to hand it over to Liman and –

LARS-JOHAN LIMAN: Thank you. Welcome to you all. We're going to delve a bit on the root of a system here this afternoon. As he said, I'm Lars-Johan Liman. I work for a company called Netnod. We're based in the small country in the far north called Sweden, and we operate one of the set of root servers. As we work through these slides, you'll see that there are 13 sets of root servers out there. That doesn't mean it's 13 machines but it's 13 identities that you can talk to.

I will hopefully later be joined by John Crain who works for ICANN and L-root but he has a conflicting commitment, so he will join us a little later. I will start on my own here.

I will walk through these slides here. We have four chapters or four sections that we will go through, and I welcome any kind of input or feedback. So if you have questions, just raise your hand and we can take it from there because it usually brings more energy and liveliness into a presentation if there's a dialogue with the audience. If I just stand here repeating what's on the slide, it's going to be fairly dull after half an hour.

So these are the various slides we're going to look at. First an overview of the Domain Name System, the DNS system. We're going to do a quick history of the root server system. How did we end up where we are right now? We're going to look at the root server system today, and how it works and what features it has. And we're also going to look at how this ties into ICANN, the connecting point between the root server operators in ICANN is called RSSAC, the Root Server System Advisory Committee, and what we're doing it right there. We're fortunate enough to have one of the two Co-Chairs of RSSAC in the room, Brad Verd in the corner. He will be able to handle any questions you have because he just took over that role from me.

So let's just dive in here. A quick overview with the DNS system. When you exchange traffic on the Internet, you have to identify the end points of the traffic. The computers that are talking to each other need know to signal somehow, "I want this packet to reach that computer over there," and the basis for that is IP

addresses. So each host or each group of hosts connected to the Internet has a unique IP address that you can send packets to, and the computer will then know how to forward that into the correct application in the computer.

Today we have to deal with two types of addresses, IPv4, which is the classic Internet addresses that we've used for a long time. If you're well-informed, you will know that we are basically running out of those but there are four billion in total and we are running very, very short on such addresses.

The good news is that there is a new generation of the Internet protocol, the Internet way of communicating on the network called IPv6 with much, much, much many more, much many more addresses. Someone said that point when it was being developed that with these many addresses, we can enumerate all the electrons on every atom in the universe several times over. So by that, I'm kind of asserting that there would be enough addresses. I'm not quite sure.

Oh, there's John. John Crain from ICANN has joined us. Good to see you.

The uniqueness of these addresses is very important. You can compare this with telephone numbers. If two people have the same telephone number, it would be very hard to tell the telephone, "I want to talk to that person with this number." So

there has to be a unique pool, a unique system of addresses and that is handled centrally by the IANA down through the Internet. The Regional Internet Registries, there are five of them right now and from these down to the various Internet service providers and down to the end users. So this is quite well-designed and well-working.

So why do we have this DNS system? We have these addresses, what's the problem? Well, their original problem was that IP addresses are hard to remember and they change. Every time you take your laptop into a new hotel or a new office or a new home, it will have a different IP address. So knowing the IP address of a host on the Internet isn't always that easy. But if you have a stable reference that doesn't change, it can point to the current address of whatever service you want to contact and talk to.

And the problem was there already with IPv4. Trying to remember an IPv6 address is totally useless. There's so many digits and letters in there that forget it, it doesn't work. We must have the DNS to help us there because the human brain handles names much better than addresses.

A modern problem is that IP addresses sometimes are shared by several computers, or that multiple addresses are used as entry point for the same service. This is used for distributing service,

creating redundancy, and load sharing stuff and the DNS is what we need to help us navigate that system to reach the right service.

The Domain Name System, it's a look-up mechanism for translating objects into other objects. Okay. So what does that mean? It means that it's a database where you can look up stuff. As with every database, you have to have something in your hand and you tell the database, "Give me something else that is related to this." If you go back to the phone book days they gave the phone book, "Here's my name. What's my number?" The same basic principle is in the system. You have a domain name. "Give me back that other information."

The other information that you're looking for is very often the IP address of something, typically a host or a service, but it can be many other things. The DNS system is not limited to addresses. You can put all kinds of interesting information in there, but you have to tell the system what you're looking for. Give me the address. Give me the email filtering information. Give me the contact person for, and so on. But the look up key, the thing that you hold in your hand when you start the look up is always a domain name.

And the same DNS system is used for all these types of information. IPv4 addresses, v6 addresses, e-mail handling, all

that is in the same DNS system. There is only one on the public Internet. You can create your own at home if you want to. I leave that as an exercise to the reader. It's globally distributed. We have the same DNS system all over the world. If I look something up here in Morocco and then I travel back to Sweden and look up the same thing, I want to have the same answer because it's one global unique system.

It also says it's loosely coherent. That means that there can be temporary differences between various points in the system, but the purpose is to always have a self-healing system where it would correct itself. If you make an update, it takes a bit of time for that update to propagate through the system. That's okay. It's supposed to be that way. But once you waited for a while, that update will have propagated, and it's once again a coherent system.

It's also proven to be very scalable. When it was invented, the number of hosts on the Internet were counted in tens of thousands. Today, we count them in billions maybe. The DNS still works. It works even better now than it did back then, but that's for different reasons. And it's also a dynamic database. You can do updates on the fly. You can send in your updates to the DNS system and have it propagate these updates very quickly.

This is the very, very typical picture of how it works. You have an end user who uses an application, needs information from DNS. The most typical example is the web browser. You type in a URL – well, okay, you use Google these days but you give the URL that contains a domain name. What the end user computer does with that is to send it off to a helping server called the caching DNS server that helps finding the data in the DNS. It's typically – it doesn't necessarily know the answer from the beginning but it will help find it. It knows how to find it. And if it doesn't know anything, it has an empty cache, it doesn't have anything stored, it will have to talk to a root name server.

On this side on the right-hand side, we have the servers that provide the information and here (blue) we have the service that consume the information. So it will talk to an entry point to the DNS system, and that's a root server, and it will ask for www.example.org. The root server will not know the answer. The root server can always give you one of two things. It can tell you "No, it doesn't exist." Not in all cases, but in some cases. If it doesn't know – if it doesn't say it doesn't exist, it will always say, "I don't know, go and talk to these guys over there." It will give a referral to some other systems. So that's the two things you can get out of the root server. It doesn't exist or go elsewhere.

In this case, the referral will be for .org, and the caching DNS service will say, "Oh, he didn't know. I have to talk to these guys

instead." The .org server will know everything about names under .org. So it will know about example.org and it will send back again a referral saying, "I don't know, you have to talk to the example people." The query will once again go to the example – the DNS servers, the Domain Name Server, for example .org who will eventually know the answer to the query that's being asked. The final answer. "This is the answer. This is what you're looking for." And that will go back and be forwarded back to the end user who has been waiting.

They send the query off to the caching resolver and then they just twiddle their thumbs, waiting for the answer, and plop there it came. And during that time, three queries has been sent off in various directions – probably more, actually – but the active parts here are the three queries sent in different directions. And this server in the middle keeps track of all of these things, so it will remember the answers for all these three questions and it will use that information if there's another query from the end user that looks for the same or similar information. Something else in .org, then it already knows this server will give me information about .org, so it can do shortcuts in this process.

This means that the root servers don't have to handle all the queries that come from end user because many of them can take advantage of cached information here. And that will relieve

the root servers of a lot of traffic, so that caching function is a very important property of the entire DNS system.

Moving right on, so the root servers don't really know much. They have a very limited set of data. I have been talking about this for more than 20 years, and about 20 years ago I would say that you could fit the root server on one of these old floppy disks 20 times over. It's not quite the case anymore because things have evolved, but it's still a very, very small data set. The thing is, that is very popular. A lot of people want to ask queries about this.

So the root servers have a small data set and it contains the list of other people that know more for various sub-domains or you call them top-level domains in this area. So it knows the list of servers for every top-level domain in the world. They get this information from the IANA, so the root servers actually define the list of active top-level domains. If it's not in the root server, it's not active. You cannot use it.

There have been some refinements, some modern additions to the DNS system. It was invented in the early 1980s. It worked basically in that shape until the year 2000. And some late additions to the system are secure DNS, where you have the ability to add cryptographic signatures to DNS data. You don't have to, but if you want to do it, there's your mechanism.

This reduces the risk of spoofing. The risk that someone else injects DNS data, and the typical place to do that is to trick this caching server into storing false data. If you play tricks here (and you can), you can kind of teach this server incorrect information, and when there are queries from end users, it will look into its cache first and say, "Oh, it's already here. Good. It happens to be false but I don't know that, so I will hand it out as the appropriate answer."

DNSSEC will help you to prevent that. But it requires that the caching resolver – going back again – this machine cannot only rely on what it receives from these servers here that give out the information here. They must also validate it, and that validation process is a mathematical process where they verify and check the cryptographic signatures. It's rather heavy mathematical process. But if you do it, you kind of remove the entire chance of someone spoofing data because the signatures will tell you that this is right or wrong.

There are privacy enhancements underway. DNS queries can actually tell people a lot. It can tell you a lot about who you are and what you're doing. If they read your DNS queries they can say, "Oh, he's going over to that website, and now he's picking up mail from that server, and now he's logging in over there. Why is that?" So if you happen to be the person operating this machine, you can look at all the incoming queries here and you

can see from which IP addresses are they coming. “Okay, this is that user? And they’re doing that. Why are they surfing over to that?” Now, if you happen to be Google, you have a gigantic server here. You can look at many millions of people just checking what they want to do.

So adding privacy here is a very important property. So there is work underway in the Internet Engineering Taskforce, ITF, which is the standardization body for Internet technology. There’s work underway there to introduce mechanisms that will give the users better privacy.

Another thing that’s been added in the past 10 or 15 years is Anycast where you can deploy a number of servers using actually exactly the same IP address, even though they are located at different sites across the entire globe, and they will provide the exact same information. So it doesn’t really matter which one you’re talking to.

I’m going to make a comparison which I believe is true for more places than Sweden. But if you think that when you dial on your phone, if you dial the emergency number – that would be 911 in the U.S. or 112 in Europe – when you dial that number, do you know whom you’re talking to? Do you know where they are? Not typically. If you dial the number in San Francisco and in Washington, do you reach the same people? I would say no you

don't. But you reach a service and the function of the service is the same. "Yes, we're here to help you. How can we help you? Do you need police, fire, ambulance? What do you need?"

So Anycast works in the same fashion but on the Internet. You send a query to that IP address requesting a type of information. And Anycast isn't used to only for DNS, it works quite well for DNS, so it's heavily used there but you can use it for other services as well. The important part is that when you contact this service that you get the same answer regardless of which service point you hit.

This is a way to improve latency and resilience in the network. It so happens that if one of these service points disappears from the network, the network itself, the components in the network will kind of shift over so the traffic goes to the other service points instead.

Root zone on the root servers. The root zone, that is the data – the DNS information that contains the list of active top-level domains. It contains the list and for each top-level domain, it also contains the names of the servers, or the next step in this process, the servers that people need to talk to or, sorry, rather the servers need to talk to, to get the information regarding that top-level domain. So they know nothing about the top-level domain except where the servers are. So it's basically a list of

top-level domains, names of the servers that serve that top-level domain and the IP addresses of those servers. That's all there is to it. Nowadays, there's actually also these crypto signatures, the DNSSEC, the secured DNS additions on there but that's only to make the information verifiable so that these caching servers can do the math and can verify that these are actually the correct servers for that or this top-level domain.

This zone file, this is basically a file on the computer. It's created and managed by ICANN, and it's under a community policy, so the ICANN community, those of us that are here discussing, we are the ones setting the policy for how this information is going to be treated. What are the rules under which changes can happen? Additions, removals, and changed in information. The policy for that is set by ICANN here. The actual part that does the machinery here, the administrative department that does the actual processing of change request is the IANA. And I think I saw Elise stepped in here so we also have the manager of IANA here.

Once that's done, once the administrative part is through, it's sent off to Verisign to be implemented in the technical system of the root servers. And from Verisign it's distributed to all the root server operators, and within each operator then distributed to the actual servers that provide this information.

So it may seem like a long chain, but it's actually not that long. The process is fairly short and quick for standard types of changes. What's taking a long time to change the policy, if there's a need for that, but personally I think it's a good idea if policy changes take a lot of time because that means that a lot of people have a chance to see them, react to them, give input, and we don't do anything in haste.

The root servers – these are the actual machines that provide this DNS information to people who send the queries. Currently, this is limited to 13 identities. Now, note that this is a carefully chosen word. There are 13 identities, and they are denoted by domain names. So you have a.rootservers.net, b.rootservers.net, c, d, e, f, g, h, I, j, k, l, n, and m. And each such identity is operated by a root server operator for purely historical reasons. Two of them are operated by the same entity, so there are 12 organizations that operate these root servers.

I work for Netnod. We operate I-root. Brad works for Verisign, so that's A and J-root. Who else was in here? Any more root sub-operators? Oh yes, John for L-root.

UNIDENTIFIED MALE: And we have Tripti on the line.

LARS-JOHAN LIMAN: Oh, we have Tripti on the line as well. So Tripti Sinha from D-root, University of Maryland.

Root servers fulfill a purely technical role. The root servers are machines who deliver responses to the DNS queries. They have nothing to do with the policy mechanisms for who gets which domain name. Once the decision is taken to put a name into service, this is the machine that does it. The machine only does what the human says and the humans who say which domain names goes where is the IANA, and they do so only under policy that's set by ICANN. So the decision system is something that happens before the root servers.

So these root servers, which are operated by the root server operators, that's a purely technical role, and in that role there is no administrative decisions taken regarding the zone content, regarding which domain names or which servers to point to in this list here. The root server operators have nothing to say about this information here, nothing whatsoever.

Moving right along. There are 12 organizations. They are very different. The main focus of all these people – and I'm happy to say that I know most of them on first name basis. We meet regularly to coordinate and make everything work. The main focus for all these people, all these guys running these servers is reliability and stability of the service. If we have to do nothing

else, the root server working, serving you, that's our primary focus. I would say that the entire rest of Netnod's business could go straight to out the door. The root server will be the last thing working at Netnod.

We also have a commitment that this should be accessible to all Internet users. We don't want to say that you are better than you, so therefore we give you better service. No. This has to be accessible to all the Internet users. The usage of the Internet shouldn't depend on access to the root servers, because it should just be there. It should just work.

As I said, we cooperate on a technical level very deeply. You will know that Brad in the corner has the telephone number that will ring the telephone next to my bed. If something goes wrong, he can reach me and so can John, and I can reach John. So these 12 organizations it's a very close knit group of people who work very well together and we can have our differences when it comes to business or decision making and so on. But when it comes to the technical cooperation of these servers, it's very close and it works really, really well.

These organizations are very diverse. Technically, each organization makes its own decision on which servers to use, which operating systems to use, which types of software to use, as long as we provide the same service towards the network.

How do we configure our routers? How do we configure our traffic filters? That's all our own decisions. But we inform the others, say, "This is what we are doing. What are you doing?" So that we can make sure that our systems don't destroy anything for ICANN systems or Verisign systems and likewise. And by having these differences, we try to make sure that if one component breaks in one system, it doesn't bring down the entire set of root name servers. If it turns out that servers from HP don't work very well – okay, we have servers from HP maybe. Okay. Our servers crashed. Verisign's didn't because they buy from someone else.

And again, if we happen to use a particular piece of software that has a severe vulnerability in it that's discovered, maybe someone can break into our systems but they can't break into ICANN's because they use a different type of software. So this redundancy or resilience in the system is important and we actually make a point out of choosing different things.

It's also organizationally very different. It spreads over companies that are government agencies. There are for-profit companies, there are non-profit companies, and there are universities. There's wide variety of types of organizations which have their different decision-making processes. We're a small company. We're 20 employees. I can step up to my CEO and say, "We need to do this. It'll cost you $5,000." And he'll say, "Okay."

Try to do that at the army research lab in the U.S. Probably it doesn't work that way, and that's good because they have a different mechanism. So if someone hijacks my CEO, that can have a very severe impact, but you can't really do that with the army research lab. So having these differences is a very good thing. It's a strength of the system.

We are also somewhat geographically spread out. You can argue either way there and say that several of these organizations are based in the U.S. but we also have organizations in Sweden, in my case, the Netherlands and Japan who are doing a good job here.

So where are we? Oh, I need to speed up a little bit. So the root server operators are not involved in policy making except sometimes as technical experts. When someone tries to do a policy that will have an impact on the system and we say, "No, that's not a good idea because it will stop working when," and our opinions when we forward them are based on the technical problems. We don't care really who gets which domain name. And the operators are not involved in modifying data. We never ever, ever change the content of the root zone and we cannot because it's signed. If we change something, the signatures will no longer be correct and every caching name server will detect that.

We are involved in careful operation and evolution of the service. We try to expand. We try to modernize. We try to follow the changing or shifting requirements from the general public on the Internet. We evaluate and deploy technical modifications when there are requests and ideas from various sides, we try to follow up there. And, as I said, stability and robustness are the very first things we look to. So if we're worried about something it's usually because we feel that, okay, that will have an impact on the robustness of the system. We don't want that.

I will be very brief on the history side here. The root server system is very old. I've been around for quite a bit, not since 1983, I'm happy to say. I was still in high school by then. But you will see here that the list of root server operator has shifted more in the old days than nowadays but there is a definitely a shift. So these are some of the servers operated in 1983 to 1986, that time period. You will see that it's mostly universities and some military agencies. You have to remember that the Internet comes from military money funding university research to create a stable and resilient network, so it made sense at that time.

In 1987 a few more were added. Still mostly universities and government agencies. In 1991, the first one outside the U.S. was added, and that's actually ours. It's listed here as nic.nordu.net. It was operated from the Royal Institute of Technology in Stockholm. I happened to be student there at that time, and

eventually grew into working with it. So six months after it was put on the network I started to work with that machine that eventually was changed into I-root. So I've been around since 1992 doing root server operations.

By '93, we hit a technical limit. There is a technical limit on how many root server operators you can fit into one DNS packet. We can have a long discussion about the relevance of that today, but in 1993 it was relevant and we have nine root name server identities which we wanted to expand on. So Bill Manning, Mark Kosters, and Paul Vixie got together and came up with an idea to rename the computers.

Now, going back, you will see that they have ordinary domain names for Internet host here and they're all very different. What they come up with here was a scheme to rename them into the same domain and in the DNS packets. That means that you can compress, you can take advantage of the algorithms and fit more identities into that 512 byte packet, so we could now fit 13 instead.

So this happened in 1995, and four new servers were established in '97 and were deployed at various places. Three of them were deployed. The fourth one remained with – regionally [inaudible] with you at Verisign I suppose – and the process kind of halted

when Jon Postel who was in charge of it all died very suddenly and sadly.

So here's how they were renamed. You will see at the bottom that nic.nordu.net was renamed to I-root and has retained that name ever since. These are the nine and four new added – they never had any different names, so they've always been just letters.

The process for adding these four was very straightforward. A need was identified. Jon Postel who ran this process, he looked at connectivity, both internal and external, so if the server was to be deployed at point X, would it be reachable? Was there a need for a server in that region? And you have to remember this was in the old days when there was only one server per letter. Was there a community consensus? Was there support in the region for this? Would there be filtering and such in front of it? So was this a neutral player on the Internet? We actually had a list of criteria which he went through when he chose and selected the four new players.

In Europe, the RIPE NCC was chosen to run K-root. In Asia the WIDE project in Japan for M-root. J-root stayed at Network Solutions which was a predecessor – which was actually acquired by Verisign later on, if I remembered correctly. Brad is

nodding. So this is now with Verisign. And L-root was transferred to ICANN as in the founding process when ICANN was created.

Jon Postel was kind of the spider in the web, so he was coordinating all the efforts when it came to root servers. But when he died, the root server operators realized that we need to continue to coordinate. We cannot just stop just because Jon died. So we had an Ad Hoc meeting where we agreed on certain principles, and we've been meeting since. We don't have a formal chairman. We meet regularly and we have meetings which are run in a fairly formal session, formal ways where we have an agenda, we know what to talk about and so on. We don't have a chairman because we don't really need that, and we don't want to appoint someone who has a higher value than the other ones. This is a ring of people that are all equals.

So these are the principles that we agreed on, to operate for the common good of the Internet. We agree that we use the IANA as the source for the data. We all need to use the same zone file. We all must use the same list of top-level domains. And we have fully agreed on that and we have fully agreed that we take that from the IANA. We agreed that we would sufficiently invest to be able to operate properly.

And if we were unable to do so, we would give proper notice and say that obviously something says that we cannot no longer do

this so we will tell the community well, well in advanced that we need to stop this service by a year from now – two years from now. And we also recognized who the operators were so we have a mutual understanding of that. Yeah, okay Verisign operates in J, and ICANN operates L, and Netnod operates I, and so on so that there's no question about who is actually responsible for which servers.

I think I'm going to hand over to John at this point so that you have a different voice to listen to.

JOHN CRAIN: This is the gadget. That's all right, I'm a little bit strange too. So let's just run you through some of the history. Ah, nothing strange, it seems to work quite well.

So this is the status of the root servers today. We're now 13 letters. The names all looked the same, root-servers.net – remember that root-servers.net. I'm going to tell you why in a minute. And these are the list of all organizations. Now, you don't need to write this down. The slides will be out there, but if you was to remember root-servers.org, that will take you to a website where you can find a lot of these information and these slides too will probably be there, I believe.

So 13 addresses but we talked about any cast, right? This idea that you can use the same IP address, the same name, the same identifiers for a service. So if you look at this map, you've got a root-servers.org – don't do it now, you're supposed to be listening to me – you can click on these, and you can zoom in, and you can see where they are. So if you see over here on Europe there's 33 servers in that part of Europe. If you zoom in on that, it will show you where they are. When you see a letter, we see D, there's just one instance in that particular location and that is D.

We call them instances, right? They are servers but we call them instances and every instance gives you exactly the same service. If I put a copy of L-root in my data center in Los Angeles and one here in Marrakech somewhere in the data center, the service is absolutely 100% identical, and that's critical, as we talked about.

So over about 500 instances constantly growing. If you feel or you worry that you're not getting fast enough service or by some miracle you've managed to set up service where you can't see a root server, that's going to be really hard unless you've got a spaceship and/or you're under the ocean somewhere. These things are very fast to get to. Come talk to us. There's no solid limit on here, we have to stop at 500. We could go to 1000, we

could go to 5000 if that's what's technically necessary. So currently looking around the 500 across the various operators.

Obviously, there's some infrastructure behind this. How does that zone file get to – Lars talked a little bit about this but here's a nice picture I get all the pretty pictures. So the TLD operators are really the ones that are in-charge of telling us what they want to have for their name servers. They're the ones operating that infrastructure so they're the only ones who know what the name servers should be called, what the IP addresses are. They talk to the IANA, the Internet Assigned Numbers Authority. And Elise and her staff are the ones who then deal with them, they go for a process and they make whatever changes are requested.

There's a process to talk between NTIA and Verisign. NTIA is the one who owns the contract at the moment. Anybody heard of something called the IANA transition? Yeah, right. So that's all about changing this little thing here, this little circular thing, that's all about changing that. So everybody communicates to make sure that we're making the change we've been requested, and everybody knows what's supposed to happen. We don't just say, "Oh, here's some changes," and throw them over the wall to the next guy and say, "Hopefully you'll get it right." There's a lot of communication here to make sure everything goes smoothly because this information is important, right? If you're a TLD you want the right IP addresses and the right name servers.

Verisign is the organization that actually generates that zone file. Remember we talked that data is in something called a zone file. It's a text file for these lists of names and numbers. And they have something they call distribution masters, and they have these in multiple locations. They don't send it to one server, they send it to multiple servers. Same information, multiple servers for redundancy because obviously you don't want to rely on one machine somewhere on the West Coast of the U.S. or wherever it is.

And then A through M, the RSOs, the root ops have machines that then go and pick up that data. We use old standard protocols here. There are protocols in moving data around in the DNS. So we do a transfer from the masters, to our own systems, and then from there – so I took the L-root – now we have a few machines that do this. They talk to Verisign. We have, I believe, over 100 machines now. We don't want them all talking to Verisign. That doesn't really make sense. So we will then distribute further, and those will go out to all the actual instances that will then actually be the ones answering the questions.

And every root server implements slightly differently for redundancy reasons, different hardware manufacturers, different algorithms for how they do it. We all talk to each other specifically to make sure that we don't do the same thing. It's

like we get together to coordinate to make sure that we'd not over-coordinated. Not everybody is buying from the same manufacturer. Not everybody's using BIND, not everybody is using whichever version of DNS server they want to use.

UNIDENTIFIED MALE: John, I'm sorry. Are you taking questions and comments now or would you like to wait until later?

JOHN CRAIN: Probably now.

UNIDENTIFIED MALE: Okay. I have a comment from [Amin] on remote participation and that there's a brand new F-root server deployed a few days ago here in Morocco.

JOHN CRAIN: Cool. I wonder if they physically came down and then left before the meeting. Actually, I think one of them is around here somewhere. I think Jim's around.

So yeah, they're always being deployed. So we know there's at least one F in Morocco. I didn't actually look or zoom into the map to see how many are here.

Once that zone is out on all those servers, as instances, that's where the resolvers go. So all the resolvers of the world can, through their caching resolvers at their ISPs, query the root servers to find out where the TLDs are. That's just a pretty picture.

So I talked about redundancy, and one of the things that you do for redundancy is, of course, diversity. And we are diverse in many ways. We're not all non-profit organizations. We're not all for-profit organizations. We have government, universities, a wide spectrum of organizations amongst the 12. We think that's healthy.

Our operational history and the kind of things we've done in the past, as well as the way we operate today is all slightly different, deliberately so. There is not one name server that all the root server operators use, deliberately so. At one point, we were all using pretty close to similar versions of BIND, and we actually went out to other people and some folks in Amsterdam that are friends of ours and said, "Could you develop something called NSD (Name Server Deamon)?" which some of us now use. So there's at least three different variance of name servers that I'm aware of on the root servers.

We, at L-root, actually run two of those variants, one primary, one secondary in case there's a problem with one of them. So

not only is there a redundancy between the various operators, we all actually do the same thing internally. We tend to run diverse equipment, diverse software through our own system. So there's layers and layers of diversity and redundancy.

What we do have is common best practices. We all follow the RFCs. We have best practices about how we talk to each other. We have systems in place that are ready to roll at any second so that we can all call each other and get on the phone bridge. I don't know how long the last one took, but it's normally a minute or so to get most of the root server operators onto bridges. So we have systems in place that we practice.

We over provision capacity. Anybody heard of DDoS Attacks, right? Everybody's had a DDoS Attacks. The root servers see DDoS Attacks. I'm not sure we're actually being attacked, but sometimes people are attacking other people and where does the traffic start? At the root. We see a lot of these tracks, and we see spikes in the data, but we all over-provision.

And of course we all should have professional and trusted staff. This is important infrastructure, and we're all very aware of that. As Lars would say, and we all know each other on a first-name basis. Most of us know each other's families. We literally can all call each other at any time. If there's one person I will always pick a phone call up from, it's one of the other root ops. I may

not pick it up from my CEO at times, but another root op calls me, I pick up. Everybody does the same thing. It's an agreement we have.

So where do we get together to cooperate? Meetings like this, mainly they tend to be operational meetings rather than policy meetings. So you may have heard of some of these organizations, the Internet Engineering Task Force, you'll see a lot of us there. RIPE, which is an operator's forum in Europe. NANOG, which is similar in the U.S.. DNS-OARC which is forum dedicated to DNS research. I actually sit on the board of that. APNIC, which is RAR for Asia and Pacific. ARIN, AFNOG and other NOGs, and sometimes we use this thing called the Internet, it's really good for communicating. So we actually use the Internet itself. Some of us can even read e-mails. Not me, I'm not very good at that. But some of these guys are.

We have this permanent infrastructure to respond to possible emergencies. That's that bridge I'm talking about. We can get everybody on voice and electronic communications very quickly and we test this on a regular basis. It's actually quite amazing to watch.

Then we do things like run exercises. You can't have an emergency response plan if you don't test it. So we test all of this

infrastructure. We test the plans about how we're going to do things.

Then we coordinate with other bodies. An important one here within ICANN is something called the RSSAC and I'll come to that in another slide. The IETF is important because that's where the standards are made, and we rely on those standards. We're part of the provisioning of the identifiers that come from those standards, so that's why we meet at IETF, a lot of people go there. We share data through that DNS-OARC organization.

At least once a year we do a full 48-hour capture of all the traffic so that researchers and analysts can go look at DNS data and learn whatever it is they want to learn. If you're network engineer, then you should remember that your DNS data is really, really interesting. Most things, if not everything that happens on the network appears somewhere in your DNS. So we're part of the research community as well in many ways.

The system has evolved in the past and it will continue to evolve. We look at all the new users and protocols. We had to add something called QuadAy records for IPv6. IPv6 uses the DNS too, so we had to study that and decide what we we're going to do about it, and how we're going to do it. DNSSEC, IDNs (Internationalized Domain Names), these are all things that in some way affect what we do operationally. Some of them, we

look at them and it really doesn't make much difference to us, and sometimes we have to actually do upgrades or add extra capacity, etc. We're involved in those things.

And we continually try to increase the robustness. We're always looking at the security and stability of the system, and the robustness of the system. That's our responsibility. We're all dedicated to doing this as a service to the community, and to the Internet, and with that comes some responsibilities. So increasing the robustness, responsiveness, and resilience is one of those.

One of those was that wide deployment of Anycast, like I said, close to 500 now. Not sure we expected it to go this far when we first started. All of us started with maybe two or three instances, and now there's some of us that have dozens and dozens of these things and it scales amazingly well.

Myths, oh I love the myths, lots of stories out there, lots of fables, lots of fairytales. Root servers do not have horns. Oh, that wasn't one of them? Yeah. Some root server operators might. They do not control Internet traffic. So sometimes people say, "Well, if everything has to go through the root server then they control the Internet traffic." No. We answer DNS queries. And guess what, no every DNS query goes to the root servers.

Caching – all right. People come to us and they want the data about the TLD, and they'll cache that for a while. Once they've got the information about where .com is, they might not come back to us for hours because they don't need to. So it's actually a very small portion of the DNS traffic. It's an important portion of the DNS traffic but it's not a large portion. The TLD see way more traffic than we do, at least the large ones. So people say that, you get all of the DNS traffic. You can see everything on the Internet. It's not true.

So sometimes you ask people which is the best root server? Which is the special root server? And there isn't one, they're all the same. You'll get exactly the same answer whether you get it from A, B, C, D, E, M, L. Whichever one you query from the perspective of, but just wanting a DNS answer, they are exactly the same. Publishing this data is completely separate from administrating data. The root server operators have nothing to do with the content. People think you control the root server, therefore you control the data.

Anybody heard of DNSSEC? Where the root server is science. In the past there was a theoretical situation where I guess you could if you were crazy going do something like this and hope nobody noticed, but you can't even do that now. With DNSSEC, we can't change the data without breaking the signature. Not that we have a word, but we physically can't now. That's a

different thing. So being a root server operator does not give you any special power over the content of the root zone.

Sometimes people say it's a bunch of hobbyists. Though these old engineers, well some of us are a little bit old – hopefully not that old – sitting in closets somewhere, running this on an old PC as a hobby. If you actually get to see some of the infrastructure of root servers around the globe, you'll see this is extremely professional and industrial-scale deployment for these systems. Five hundred systems coordinating globally around the globe between 12 organizations that coordinate. It's not a hobby. Definitely not a fun hobby if it was one.

There are more than 13 root servers. People always say, "Well, there are 13 root servers, and they mean there are 13 root server identities." There are hundreds and hundreds of instances. So from the technical perspective of getting a query, there's hundreds of root servers.

Sometimes you hear that this is all separate and uncoordinated and these are a bunch of separate things. But we do coordinate, but only about those things that we talked about. Coordinate enough to make sure that we are providing that service to the community, and that service is quite simply answering DNS queries about TLDs.

UNIDENTIFIED MALE: John, I have another question.

JOHN CRAIN: Yeah, go for it.

UNIDENTIFIED FEMALE: John, we have a question in the chat room from Said Zazai representing NITPAA in Afghanistan. The question is what sort of support do you or the operators can provide in terms of installing new instances in countries where government and private sector have no interest in financing it?

JOHN CRAIN: It's different for each root server. We all do this differently. So I can talk to L-root. What we've done is try to go for a very small profile box that is relatively affordable, and I understand that affordability is relative, and then we manage all the operations. So that's one way of doing it. Some, I believe, will actually do some of the funding of the hardware.

So the real thing to do is to come talk to one of us, and I don't think it really matters which one. Maybe come to Lars here because he has nothing to do at the moment and talk to him and find out if they can help and if not, if they know somebody who can, because we do coordinate.

If you're running an ISP, getting a box for this should not be outside your margins. The operations cost – you plug it into your network, and then in our case we do all the operations, and I think in every case the operator does all the operations, so they're not prohibitively expensive. But I do understand that expensive is a relative thing. I don't know what our boxes are at the moment, but we're actually looking at doing an even smaller cheaper box. So come talk to us, we'll figure it out.

UNIDENTIFIED MALE: Do you want to break for questions on this section before we move to the RSSAC stuff?

JOHN CRAIN: Absolutely.

UNIDENTIFIED MALE: I have a question. So the root zone is a list of TLD and a list of servers for each TLD, so the TLD don't change very often, right? They change, but not a lot. So I assume if you get traffic and you have to expand the service, the name of the servers change a lot. Is that the reason why you have to serve? Otherwise, the caching would just in the end stabilize to not querying the root server a lot.

JOHN CRAIN: Caching has something like a time limit on it, so you don't cache forever because things change. All right. So there's always constant discussions about how long the cache should be. And things do change, so you can't rely on that. I don't know how often they change, but I'm sure Elise – because I don't want to put her on the spot – has a better idea of the changes because they actually manage them.

ELISE GERICH: So was the question, how often do the root server name servers addresses change?

JOHN CRAIN: No. How often does the root zone change?

ELISE GERICH: The root zone, thank you. The root zone is compiled twice daily, and the changes in that are dependent obviously on the various TLDs that submit request. And over the last quarter we've had over a thousand requests for root zone changes. So that's twice daily, but it depends on how many come in at any point in time, and a thousand were done over the last quarter.

LARS-JOHAN LIMAN: If I can add to the caching situation, you're quite right about how it works but the thing that a lot of people don't think of is that the root servers also get a lot queries for things that do not exist, people that mistype stuff, people that cut and paste the wrong thing. So we have these massive amounts of queries about things that aren't even top-level domains, and you cannot cache that. So that is something that drives the query rate for the root servers and the general growth of the Internet.

JOHN CRAIN: Yeah. As you can see, everybody thinks it doesn't change very often but when you've got a couple of thousand TLDs in there, it pretty much every time we update this like to be some kind of change in there.

UNIDENTIFIED MALE: Any other questions in the room right now? I have one more remote, but I wanted to give in the room a chance. Okay, we have one more remote.

UNIDENTIFIED FEMALE: John, one more question from Jahangir Hussain representing a Technical University in Bangladesh. The question is, is it possible to share some information about the daily or monthly traffic of the root server?

JOHN CRAIN: Yes. Most of the root server operators actually publish statistics. If you go to root-servers.org, you see links for each of them and you actually see statistics. Some are live, some have a 50-minute delay but the statistics are out there, absolutely. And then there are studies that are done about this as well through OARC, so there's lots of data out there.

For instance, if he wants this information related to doing an install and actually getting one, then he should probably just come talk to one of the operators. Because many of these instances don't do much traffic. Now, people think these things get gigabytes of queries a day but it's not. Across the whole system of L-root – actually I don't know. Do you know what I is doing at the moment, roughly, in packets per second?

Okay. We can find out. These data is all out there publicly, but we're talking megabytes, not gigabytes across a letter. So if you get to one of these instances, you just divide that down basically. So some instances will get a few 100,000, and some may get like a low megabytes but none of them are massive infrastructure drains but they're all over-provisioned.

Has anybody heard of the Root Server System Advisory Committee or RSSAC? This is a ICANN thing. So ICANN has a lot of advisory committees, and obviously the Root Server System is

probably going to be important to people here at ICANN because a lot of you are interested in this DNS thing. So there's an advisory committee which has been around for a long time, I think as long as ICANN has.

It has a very specific role, it's to advice the ICANN community and Board on matters relating to the operation, administration, security, and integrity of the Internet's root server system. That's verbatim from our role and remit documents, and it's pretty much always been exactly that since day one. So basically, we're an advisory committee. We advise the board and the community on anything that may affect the root server system, which is what I just said. But we also provide advice if asked by other bodies within ICANN, as well a part of the community and outside. We look at issues and advise on them. And the Board can ask us question if they want to and we will try and answer it.

The root server operators are represented inside RSSAC, but RSSAC is not involved in the operations. RSSAC is an advisory committee, it's not an operational committee. We have a horrible picture here, but down here – you can see the red box, but you probably can't see what's in there. It says Root Server System Advisory Committee. Above that you it's the Internet Engineering Task Force, Security and Stability Advisory Committee, these are Board of Director positions. One of the things that RSSAC does is it appoints a liaison to the ICANN

Board. Currently, into that is a lady called Suzanne Woolf. She is around somewhere and you should feel free to ask her questions. So as you can see we're just part of the system, and one of the advisory committees.

So who's in RSSAC? Well, the root server operators get to appoint an individual and an alternative because the phone calls for these things are often at horrible times, and we don't want the same person getting up all the time or one may just not be available. And then we have what we call liaisons and I'll get to those in a second. The other thing we have is something called the RSSAC caucus. These are volunteers and subject matter experts, and they get appointed by RSSAC, and I'm going to go into more detail on these.

So we have two lovely co-chairs, Brad I believe is in the room somewhere. He's hiding over in the corner, but now you know what he looks like, so hunt him down if you have questions. Tripti is on the phone, so you can't hunt her down but she often comes to ICANN meetings so if you see her around, if you come to a future ICANN meeting, say hello to her. So Brad is A and J and Tripti is D. We refer to each other by our letters sometimes. It's very sad I know, but we do.

So who are the liaisons on the committee? The IANA Functions operator, the people who run the IANA thing, specifically in this

case, Elise. The root zone maintainer, the people who do that maintenance of the root zone, which is Verisign. So Verisign has one for their root servers, like a person on the committee, and then they have a liaison for this specific purpose. The IANA Functions Administrator – that's the NTIA, or that IANA transition thing, remember? Most of this week will be about that.

The IAB, the Internet Architecture Board. That's the lead committee for the IETF, if you like, they also get to send us a liaison. SSAC, the Security and Stability Advisory Committee of ICANN. They also have a liaison. We call these incoming liaisons, because those are people sending the liaisons to us, and then ICANN gets to send out a couple, the ICANN Board, Suzanne Woolf and the ICANN NomCom.

So the ICANN NomCom is a body that selects board members, and we send somebody to that as well. All the committees do that, it's just part of the ICANN process.

But forget the committee, that's not where the work happens. The real work happens in the caucus. So the caucus at the moment has 71 technical experts because that's what most of them are, 45% who are not from the root server operators. When we're doing work which is mainly writing documents and advice, we look to the caucus to help us do that. We're a reasonable-sized committee, but we don't have time to deal with all these

issues. Plus we want to get a more diverse group of people advising us. So there'll be public statements of interest, if you want to join.

Every member will have why they want to join. They have to say, "I want to join the RSSAC caucus because they have good coffee and biscuits," or whatever their excuse is. It's normally people who are very interested in protocols, and technology, and operations. If they actually do work, i.e. they write documents or they're co-authors of documents, they actually get the public credit. You get your name on that document, so it's good to join the caucus.

So the idea of the caucus when we founded this was really to expand that pool of expertise and to give us more resources so that we could do more work. It was clear that we couldn't do it all on our own. The other thing was we wanted to be very transparent about what RSSAC does, so pretty much everything is out in the public. That's why we have that public statement of interest. You could see who's on the caucus, and why they're interested, and what they do for work.

And we have a framework for getting things done. We form into working groups. If we have a topic, it will go out to the caucus mailing list and we'll say, "Hey, here's an issue," and somebody from the caucus can do the same thing. "Is anybody interested

in working on it?" And then they'll form the working group and off they go and magically at the end, hopefully, advice comes out.

This is where you all get to apply after this session. It's RSSAC-membership@icann.org. We have a membership committee that then vet people to make sure that they're not really just coming for the coffee and biscuits. But if you have expertise, and you think you'd like to be involved in these things, please, please, please apply. We can use more experts always.

And these are kind of things they publish. So in good old ICANN fashion, we do the committee name and then a number. We've not got many of these yet but we're going to have hundreds of them, and they're going to be the most widespread, best documents on the Internet. It's so exciting, or not.

So 001 is about setting the expectations for service from root servers. We've had some documents in the past in RFCs, etcetera, that talked about capacity and some of the things you're expected to do. But we decided it made sense to have one as an RSSAC document. And it's actually published in tandem with an RFC. So it doesn't just stand on its own. We work with the IETF to make sure that we covered everything we needed between the two documents.

002 is on measurements of the root server system, and we all know we need to measure these stuff. There were questions raised in various points along the new TLD processes about what happens when you distribute new TLDs, what does it do to the system? As you make change in the system, how are we measuring the system to make sure there's no breakage, or that we understand what we have to do for capacity. All these standard engineering things that you should be doing. So there's a document, RSSAC 002, that specifies some of these right down to the language you use to actually send the data. It's pretty geeky.

003 was interesting because we had a whole discussion about those time to live, those caching periods, of root zone data. So a group went off and did a report – and I think we eventually decided not to change anything but we did the research first. So that's what we research about whether or not we should shorten or lengthen the time to live of root zone data. I think in the end it came out to you could but it's not really worth it. So at the moment, I don't think there's any change happening there.

We have some statements, which are different to reports. We did a workshop last year, and I think that's the first time that we all got together, and locked ourselves in a room for two or three days and try to fix some issues out. For transparency, we issued

a report from that workshop. So if you want to see the kind of things we talk about, go ahead.

We did comment on the ICG proposal, part of the IANA transition stuff. We're asked to comment, so we did. And the CCWG work stream reports, we did those. I'm not sure where the IAB liaison document, statement is.

LARS-JOHAN LIMAN: That's actually a common understanding between the IAB and the RSSAC about why we have this exchange of relationship there, and what the different roles are in respect to each other. So it's a very short document. It's just to make sure we all know why it's there.

JOHN CRAIN: Okay. It just says we like each other. Excellent. Current work or ongoing work. Believe it or not, there is no good documentation of the history of the root servers or there wasn't. So we worked with some of the staff at ICANN and the RSSAC committee to try and document some of these things. Like some of these data that Lars were showing came out of that work. We knew the information, we just hadn't documented it. So that document is about to be published, and we will get another slide in a second.

It's actually a fascinating read because even I didn't know some of these stuff. You know different people have been around for different lengths of time and were viewing the changes from different perspectives, so it's a really good read. I'm hoping that it'll be out soon.

So RSSAC 002, which is the measurement thing version 3. Some of these documents are living document, whether or not we should. Simple technical things. We're measuring the size of the zone. Is that interesting or not? These are some of the questions they have. How to measure and report load-time metrics? How long does it take the zone to propagate from the start to the end and then to load into the system? The very exciting stuff, maybe.

A lot of clearing of ambiguities in codes. If you're not in this world, these probably don't mean much to you, but if you get involved, you have different documents, have different things. This one is really about how we measure the effect on the root servers and the system when we change things on the outside or on the inside like new TLDs.

So the naming scheme. This is about the naming of the instances, not about root-server.net but how do we all name the instances? So at L-root, we use airport codes. So if you go and use the DNS to look at which L-root you're looking at, you'll probably see it's got an airport code in the name. Yeah. That's

interesting if you want to know. There are identifiers in there as well that tell you where it is, but it's probably a good idea to document how we're been doing that in the past, and how we're changing that. And then we'll see whether or not we should have a uniform mechanism or not. I don't want to say what the outcome of the committee is, or the working group because it's not done. This kind of work is the stuff that the committees do.

So yeah the working groups work on various documents. And if you're in the caucus, or even if you're not you could always send an e-mail to us that can say, "Hey, I think I see this problem, can you work on it?"

Some resources here from the committee. This is an ICANN committee, so they're all under ICANN.org. And most importantly, the membership for where you want to request to join the caucus if you're into DNS and things, and please do. Any questions about what the RSSAC is up to, or any of the documents if anybody's ever read any of them or anything else?

LARS-JOHAN LIMAN: Can I jump in? As you think about your questions, you asked about the stats for I-root. Yeah, and I've misplaced them again. Oh here. So I-root, which is a cluster of roughly 55 machines across the globe is right now operating at 55,000 queries per second, but that's for the entire cluster of all machines

combined. But it's a very large difference between the one with the most heavy load and one with the most light load. So the heaviest one right now is in London at close to 6,000 queries per second, and the lightest one is Vanuatu which is 1.2 queries per second without thousands.

JOHN CRAIN: Yeah. That matches pretty much what L-root has as well. We have the big ones that are up there and then we have the one or one point something queries per second in a few places, and that's fine because those queries are very important to the people asking them.

LARS-JOHAN LIMAN: We wouldn't have a site in Vanuatu if it wasn't important. And actually our site there kept on working at the earthquake – was that a year ago? Roughly. And kept the local communication within the nation working. Even though we couldn't reach the site, it's still was operating isolated from the rest of the world and providing service until it got reconnected again and we could continue to update. So it continued to work even though the entire nation was disconnected from the world.

JOHN CRAIN: Okay. Do we see any questions?

UNIDENTIFIED MALE: You made everyone a professional now.

JOHN CRAIN: Well, you fill out the questionnaire on the way out.

LARS-JOHAN LIMAN: Then you come and talk to us in the corridors. We love that.

UNIDENTIFIED MALE: I want to take a moment and thank John and Liman for doing this. It's really important stuff that the root ops do and I appreciate you guys talking to everybody about it, so thank you.

UNIDENTIFIED MALE: And with that, that wraps up Day 1 at How it Works session. So if you fell asleep, it's okay, we got the complete repeat of this tomorrow. Actually, the gentleman who's sitting up here, who was talking to John, he is the Managing Director of Technology for the W3C, the World Wide Web Consortium. And he had the misfortune of having the first meeting, the first session on the first day on a Sunday here, so we only had, unfortunately, five people in the room and look at all of you. So while you guys were sleeping and are enjoying breakfast or at a meeting or something, we didn't have anyone here.

So if you want to know more about the World Wide Web Consortium, it's actually a little bit later tomorrow. It's after the opening ceremony, and then we'll have it in this room again. You guys can find out all about the W3C, and I encourage you. It's actually quite an interesting presentation. So if you've got some cycles to spend, spend them with us.

We did send you a handout, a questionnaire, we did run out. You guys overwhelmed us with what we thought would be the attendance so I apologize for those who didn't get it but I'm just going to mark that you loved it if that's okay. Please take the time to fill it out. We want your feedback. We listen to it. We'll read it. We're looking for things to do that's of interest to the community. So it takes two seconds to fill it out, and we would appreciate it. And with that I'm going to say thank you all. Hopefully we'll see some of you back here tomorrow at 10:30. 10:30 is the first session. Otherwise, we'll see you around this week. Thank you.

One more thing is that this presentation we posted on the meeting website. It's processed, so it's going to take a day or two to get there.

UNIDENTIFIED MALE: So that'll be under dnsrootserver.org.

UNIDENTIFIED MALE: Exactly. It will be there too, but if you go to meetings.icann.org, in the agenda, you can drill down to the session and it'll be posted there.

UNIDENTIFIED MALE: But you will also post it to the site?

UNIDENTIFIED MALE: They said that they intended to, yes.

UNIDENTIFIED MALE: Okay, cool. Thank you so much.

[END OF TRANSCRIPTION]