DUBLIN – How It Works: Root Server Operations
Monday, October 19, 2015 – 15:45 to 17:15 IST
ICANN54 | Dublin, Ireland

LARS LIMAN: Today is to go through the overview of the Domain Name System; as such, the technology, the DNS system that we have, to look a bit at the history of the root servers, how have they evolved over the years, the root server system today, its features, and also some recent RSSAC activities.

RSSAC is the Root Sever System Advisory Committee. It's one of the ICANN bodies. It's an advisory committee to the Board. I am happy to serve as one of the two co-chairs of RSSAC. We are two people sharing the load. It's a very small advisory committee, and we have a very, very narrow focus. But we'll get into that later on.

There we are. Next slide, please. First, a bit of overview of the Domain Name System. Next slide again.

The Internet has computers connected to it, and the fundamental identifier for computers on the Internet is the IP address, the Internet Protocol Address. Each host that's connected to the public Internet needs to have a unique IP address in order to be reachable. You have to be reachable even

if you only work as a client doing surfing on the network because there are packets going back and forth; packets sending requests, packets bringing web pages back, for instance. Or telephone voiceover Internet communication. You talk both ways. There are packets going both ways. So each of these end points needs to have its unique IP address for the network to be able to locate the exact host.

Sometimes, you group a group of hosts behind a single unique address on the public Internet, but then you have unique addresses in your local environment. That could your home, or that could be your office.

Today we see two variants of IP addresses. We have the traditional IP version 4 addresses, which are denoted by four octets, four numbers between 0 and 255. And we have the more recent IP version 6 addresses, which are much, much, longer and they use hexadecimal digits. You will see the normal digits, plus the letters A through F.

The uniqueness here is guaranteed by having the addresses allocated from a single pool of addresses. So the IANA, which is a function that is operated by ICANN, has the overall responsibility for the total pool of all addresses on the Internet. But it would be rather burdensome for IANA to hand out IP addresses to each and every user on the network. So it's a hierarchical system

where IANA delegates large sub-pools of addresses to regional Internet registries (RIRs) in the five regions: Europe, Africa, the Asia-Pacific region, Northern America, and Latin America.

Within each of these regions, there is again a next-step in the hierarchy where the region delegates addresses – or again, sub-pools of their pool of addresses – to the Internet Service Providers, which in turn then delegate addresses to the users. Next slide, please.

In the beginning, the IP addresses were actually quite short. They were just one byte in the very, very, very, very early days of the Internet. But this IP version 4 form that we see today has been around for a long time.

One of the problems is that they're rather hard to remember. The human brain is much better at remembering names than numbers. Also, the IP addresses also change. You move your computer today to a new place and then get a new IP address. How do you tell your friend that? Or how do you tell your friend that your Internet Service Provider decided to assign a new number to your computers?

So that's something that prompted the use of names instead – names that could be mapped, that could be translated into these numbers.

We also have the modem problem, where an IP address may be shared by a group of people or a group of computers, or you have multiple IP addresses that service one single entry point to a particular service. For instance, if you have servers located in different geographical areas, you may want to promote them using the same name, or you may want to have a resilient system where the client can choose any one of the addresses and you still want to promote it under one name. Next slide, please.

The Domain Name System is a hierarchical system of names. Now, this picture I've shown many times. It can very easily confuse you if you don't look at it very carefully. What I'm showing here is not a network of computers. These are not computers connected with links. These are how the names in the DNS system are tied together. So this has nothing to do with the host on the Internet and fibers and cables and what have you. This is a mathematical picture of the database; how the names are tied together.

There is a hierarchy in the names, where you have a common root at the top. This is a root, not a root server. This is the root in the name system. Underneath, you have a first layer, which is commonly referred to as the top-level domains. Then underneath you have the second-level domains and so on.

These layers you connect together to form a domain name. So the name down here is expressed as www.cmu.edu., darpa.mil, alpha.usmc.mil, mil.uk. So this is how you read the names.

Then you have a relationship between the people or organizations that are responsible for administrating a part of the namespace. So this is the Domain Name System namespace.

For each of these names in the namespace, you can map other types of information to these names. The DNS system is a big database, and you can do lookups in the database. The thing you look up is the domain name. But you can connect different types of information to the names. The most typical one is the IP address. I have what I consider to be the name of a computer. I ask the DNS system database, "What's the IP address that is connected to or mapped to my name?" The database system will respond with the address. But that's far from the only thing you can put in there.

For instance, if you use e-mail, on the right-had side of the @ sign, you have a domain name. It's not the name of a host. It's the name of a mail domain. You use the DNS to find, "Which server should I talk to when I want to send mail to this domain?"

So these are just two small examples of what you can put in there.

The DNS system is globally distributed. Mind you, when I talk about this, I talk about DNS on the public Internet. So the DNS system on the public Internet is globally distributed. There are DNS servers in each and every corner of this world, serving its little, little, tiny bit of the big puzzle.

The system as whole is not coherent. It doesn't have a well-defined state at any given moment. There are always updates happening. There are always reloads happening. There are always small changes happening all over the place.

But overall, it's rather consistent. It works. It fits together. We say that it is loosely coherent. It kind of works because there are margins and there's leeway in all the corners. But it is dynamic. It changes all the time, and there's no way you can take a picture of the entire DNS system as it looks right now because it will change a millisecond later in Japan, and then again in Argentina, and then somewhere in South Africa. So that's how it looks. Next slide, please.

When you want to a lookup in this database, there are two sides to it. You have a side that provides information. That's where you store your DNS records that you want other people to be able to look at or look up. This is where you publish your information. And you have the client side, where you consume

the information. These are all the users who want to have access to the DNS information on the network.

In the middle is something that has a lot of names. One of the names is caching server, so let's use that. This is a helper's function, which is located somewhat close to the end user. This could be at the ISP. It could be a large office which probably has one of these caching name servers. But it's on the client side of the big network. So here in between is the big Internet.

What happens when an end user wants to find out the IP address for the web server wwww.example.org? It will ask for a specific record type in the database. It's called an A-record (Address Record). It will send that DNS query to the caching DNS server.

Now, caching here means it has a cache, which is a database concept of a temporary memory for things that it stores for a shorter period of time. So if it starts out with an empty cache, it knows virtually nothing. It receives the query and it will have to try to find the answer because this caching DNS server works on behalf of the end user machine.

It will then go to a root name server to try to locate this record, because when it's empty, it believes that the root name server knows everything. Turns out that the root name server doesn't know everything. It knows only a very, very, very, very tiny

fraction of all the DNS information. But this guy doesn't know that yet.

He's quickly informed because when he sends the query to the root name server, the root name server says, "I have no idea. The only thing I can see is that the name you're asking for ends in .org, and I happen to know where the servers for .org are located. I can tell you that." So that's the response that goes back here.

This machine realizes that it hasn't reached the final answer. It only got a referral, a step on the way. So it has to issue the next question, going to the servers for .org. Again, it hopes that that server will know the final answer. So it asks exactly the same question again: "www.example.org. What's the address?"

This machine will say, "I have no idea. But I can see that the name you're asking for ends in example.org, and I happen to know where the name servers for example.org are located on the network. So here's the list."

Okay. One more time. Now the caching DNS server talks to a name server for example.org, and it still hopes that it will know the final answer. "What's the A-record for www.example.org?" This machine happens to know. It will provide the answer, the record, that the caching DNS server is looking for, and that server will now eventually provide it back to the end user, who

can then use that IP address we were looking for to connect to the web server at that address to retrieve the webpage you're looking for.

So the root servers are the entry points to the system. You can always ask a root name server, and it will never know the answer. But it will always give you a referral. It can always give you one of two answers. One is, "I see that your question ends in something here .org (.se for Sweden, .cl for Chile, .museum, .stockholm, .paris – whatever) and it will know the list of servers working one level down, and it will respond with that. That's one type of answer it can give.

The second time of answer is, "That name does not exist." It will not know whether a name that ends in .org exists or not. It can see that it ends in .org, and so far that's okay. So we'll have to come to the next level until we know [inaudible] don't exist. But if it ends in a top-level domain that doesn't exist, it can say so right from the start. If you ask for [inaudible], it will say, "Does not exist. Go away."

UNIDENTIFIED MALE: [inaudible]

STEVE CONTE: No. So this is how it all works. Now, there are two things. The caching DNS server here will hang on to every one of these responses – even only the partial ones, the referrals – because that will help the caching name server to take shortcuts.

The next time another end user comes and asks for test.org, it already knows that there's no point in talking to the root name server because it will only give me a referral to the org servers. It doesn't know anything about things under .org. So I can skip that step. I can talk directly to the org servers, and the start the traversal going down from that point.

This saves a lot of queries because every query where the caching server already knows something about that specific top-level domain, that query does not have to go to a root server, and the root servers don't have to bother with that question. So that relieves a lot of pressure from the root servers. Of course, that works on every step on the ladder here. So the more you use one of these machines, the more it learns, and the quicker shortcuts it can take in the system.

Now, there are also mechanisms in there to make sure that you don't have information that's very old in here, so eventually the cache will kind of die, and things will be ripped out. But it will give a lot of good shortcuts here and relieve a lot of pressure from the top of system.

The third thing is that the DNS resolution looking up this information is something that precedes the actual transaction you want to do. None of you in here as Internet users are interested in the records in the DNS database. Really. Not when you use your computers. You're interested in making a web surf, making a phone call, doing a bank transaction – something else. That's what you want to do, and the DNS is helping you locate the service that you want to talk to.

So this is something you do before you do your real transaction, and you terminate the DNS resolution process before you do your real stuff. First, DNS lookup, then web transaction. First, DNS lookup, then phone call, or what have you. Next slide, please.

I think we went through this. The root servers know only who you need to ask next. It knows for each and every one of all top-level domains the list of servers that you need to talk to to be able to take the next step down in this tree. The root servers truly have an updated list of all functioning top-level domains. And we talked about the caching as well. Next slide, please.

There have been some modern refinements to the DNS system. Three are mentioned here. DNSSEC was gradually added to the DNS system, and this is something I'm very glad to say. DNSSEC, the Domain Name System Security Extensions to DNS, add

cryptographic signatures to the DNS data. That gives the user – or most often, this caching DNS server – the ability to verify the records.

You have to remember that the records have traveled from, in the previous picture, the publishing side. They travel across the public Internet and then to the caching DNS server. That journey across the Internet, what happened with the packet during that journey? Did someone modify it? It's been passing at least ten routers, and probably tons of links, and we have no idea what happened.

But the cryptosignatures will give the caching DNS server the ability to verify the content and see, "Okay, this is actually what the publisher meant. This is what he said. It hasn't been modified in transit on the network."

This reduces the risk of spoofing, but it also means that the client side actually has to go through this process of validating, of checking the signatures. That's a lot of mathematics, so it takes some resources. But it substantially raises the security bar to something much better.

To be honest, I didn't think that this would be implemented in the large DNS system during my lifetime, and that was not for technical reasons. That was for political reasons. But I am very pleased to say that the root servers operate with the signed

zones since several years back, and that is actually very much thanks to the U.S. Department of Commerce.

There's work going on with privacy enhancements. When you send your DNS queries across the network, you reveal what you're looking for. These packets are clear text packets, so everyone who is able to snoop, to look at the packet, can read what's in there, and they can read what you query for. By doing that, they may be able to deduce some information that you don't want to reveal.

So there's work going to enhance the privacy in the queries in order to make it more difficult for people to draw conclusions from the queries that you send. There are various paths being explored – encryption or just sending parts of the queries and so on. But there is work going on.

Another interesting thing that we introduced some 12-13 years ago, explicitly in the root server system, was Anycast, where we have multiple servers spread all over the globe providing DNS service from the same IP address.

Now, I told you already on the first slide that every host on the Internet needs to have a unique IP address. That was only a modified truth. If you know what you're doing, you can provide service from the same IP address on multiple places on the Internet, but it's something that requires very careful

coordination and synchronization between these machines so that you provide exactly the identical service from all of them.

But if you do that, you can improve on the latency and resilience of the system, because with Anycast, you can remove any one of these servers for an extended period of time, and no one on the network would even notice. You can put it back again in service and it will start to carry its own load, and you can move another one for service or upgrades. By putting the servers closer to the end user, you can make the latency – the response time between the user and the service – much quicker. Next slide, please.

The root servers versus the root zone. These are two concepts that we need to distinguish. The root servers are machines. They provide the service of responding to queries. They receive the DNS queries from these caching DNS servers, and they respond with the appropriate records of the appropriate response in the DNS protocol back.

It's currently limited to 13 names. We can have only a list of 13 names in there, but each of these names can map to many hosts using the Anycast technology. It may show up as a single IP address, but in the case of I-Root, which I happen to operate and work for, [we're] above 50 sites that provide the service using the same IP address. I would say most of the other root server operators also provide this Anycast service, so we are up in

several hundred sites across the globe, which provide root servers.

The root server is a purely technical role. The role is to serve the contents of the root zone. The root zone is the content of the database, and this is the machine that provides that content to the Internet.

The root servers are operated by the root server operators. There are 12 organizations that operate root servers, and we will get into that later on.

The root zone is essentially the list of top-level domains, and for each of these TLDs, it's the name of those servers that provide service for that specific top-level domain, the next step in the hierarchy when you follow the namespace going downwards.

Mind you, the root zone makes no distinction whatsoever between gTLDs and ccTLDs. We are now talking about the technical side of stuff. The distinction between ccTLDs and gTLDs is purely a political one, or a policy-related one. It has no impact whatsoever on the root servers. Both the root server operators and the content of the zone makes no distinction between these two.

The root zone is created and managed by ICANN. When it makes changes to the root zone, it follows the policy that's set by the

community. Here this distinction kind of creeps in because if someone wants to make a change to a ccTLD, that's a specific set of policies, and if you want to make a change to a gTLD, that's another set of policies. But that's for the process of changing the content of the zone. Once it's been changed, it goes into the root servers and they treat them all equally.

It's compiled and distributed by the company VeriSign to all root server operators, and all root server operators pick up exactly the same zone file from exactly the points we've agreed on with VeriSign. This is also in itself a very resilient system, so there is not a single point where we picked up, but there a number of points which are well synchronized by VeriSign where I as a root server operator can pick up the zone file.

If it goes really bad somewhere, I can pick it up from a normal file server. I don't even have to use the DNS protocol. And to be quite honest, so can you. The root zone is fully public. You can download it on your laptop here and now and have a look at exactly the data that we show. There's no difference. It's the zone file. So there's no hidden magic there. Next slide, please. T

here are 12 different organizations that operate root servers. The reason for 12 versus 13 is that one organization operates two of the names. These organizations have a very long history doing this, focusing really hard on reliability and stability. There four

items here are truly the key issues driving the root server operators. So reliability and stability is one of them. Accessibility to all Internet users. We honestly work very hard to make sure that all users on the Internet can access this information. We work together. We cooperate on technical matters, and that works very well. We have a long history. We know each other well. We know what we need to do, and we meet regularly. Three times per year, we have technical coordination meetings where we sit down and hash out the things that we need to talk about, which are usually not that many on the technical side of it.

We try to do this in as much as a professional fashion as we can because this is very, very important to us. If any root server operator made a mistake, the entire Internet will know about it. That's something I don't want to have on my business card, let me tell you.

These organizations are very diverse. It ranges from organizations that are government-funded and operated, to private companies, to non-profit organizations, to academic sites, universities. So we have a wide variety of types of organizations. That's actually a good thing because if there is a problem that affects one type of organization, the other ones will not be affected by that, and we have a resilience between the different types of organizations.

It also goes technically because each and every one of us chooses for his own organization which technology to use, which types of computers, which operating system, which software. Everything we choose for ourselves.

And we don't choose the same. Sometimes we talk about it, and sometimes we even talk about it from this standpoint so that we avoid choosing the same. "Oh, you three guys all run this specific software. Then I'll choose something us so that we aren't vulnerable to the same problems with the software."

We are also spread geographically to some extent over the globe, not only with all the servers, because that's really widely spread, but also the home bases, the headquarters, for the organizations are somewhat spread over the globe, which helps because we're in different jurisdictions and so on. Next slide, please.

So what do we not do? The root server operators are not involved in policy-making. We hate that. There are so many people that are so much better than we are. We like the computers. We like the network. We like to make sure that everyone gets the response. But the policy for who gets what name? I don't really care, to be honest. There are so many people out here at ICANN who care about that that that slot's filled, so I don't need to worry about that.

Another thing that we do not do – and we are extremely careful here – is to modify the data. The data we work with is the root zone. We are extremely careful to copy exactly what we're given and to make sure that that's exactly what's provided to all the Internet users without any modification whatsoever. If you have the process of publishing a book, you have an author who writes the book, and you have an editor who may make a few changes or suggested changes to the book. But then you have someone who actually prints the book. This is the big machine that prints the paper and binds it into books. That's us. We have no say about the text in the book. We print the pages. We give it to the readers. That's what we do. So we're the publishers. We're not the authors or editors.

This is actually where DNSSEC kicks in. The root server operators do not have access to the necessary keys for modifying the signatures in DNSSEC, which means that if any one of the root server operators tried to modify some information in the root zone, the consumers, the clients, would immediately realize that because the signatures are no longer correct because the root server operators don't have access to the necessary key to generate new signatures.

The operators are involved in care for operation and evolution of the servers. We need to follow what's going on in technology. There are new standards and stuff coming out. There is

expansion of the network. There are new functionalities – IP version 6, DNSSEC, and what have you. We need to follow that. We need to value to deploy suggested new tech modifications to the protocol, to the servers, and to the network because we do operate some parts of the networks near our servers. We work as [inaudible] piece when we provide the service from our Anycast nodes.

But the main focus for us all the time is the stability. In my personal view, I would rather provide a service which is slightly slower, slightly older, but works every day in the week than to be at the leading edge, providing the most modern thing, and it stops working for two hours. I wouldn't be a happy provider.

So we take it very carefully. We modify slowly. We evolve with time, but we are definitely involved in finding the new standards, finding the new protocols and so on. So we're aware, but the stability and resilience is our first and foremost priority. Next slide, please.

I'll go slightly quickly through this: the history of the root server system. Next slide, please.

Back in the 1980s, the hosts on the Internet were listed in a text file. This was like a Word document with name of the host, IP address, name of the host, IP address, name of the host, IP address. It grew longer and longer and longer and longer and

longer. This file had to be copied to each and every machine on the Internet. Eventually, that became too much of a hassle.

So a few smart people got together and invented the DNS system. That called for root name servers. Just among them, more or less, they figured out can you run one of these? Do we have enough spare capacity on your computer? Are you well-connected to the Internet? Yes. Sure. So there were four servers on the Internet in the early days. We're now back in the early 1980s. You haven't even heard anyone mention the Internet back then.

You can see here where they were located. They were located at the Software Research Institute's Information Sciences at USC and the Ballistic Research Laboratory. And the different types of software they were using. JEEVES is no longer with us. That type of software has died out. BIND actually is still used by several root server operators. Next slide, please.

Eventually the system was expanded back in 1987. So the two ISI servers were combined into one, so there's one less in the upper half here. But four new were added. You can see again that these are mostly military and research or academic institutes. We have the University of Maryland here. We have the U.S. Air Force. We have NASA.

If we continue to the nest slide, in 1991, the first one outside the U.S. was added, and that's actually our server in Stockholm, Sweden, at NORDUnet which is the research network that covers all of the Nordic countries, being Sweden, Finland, Norway, Denmark, and Iceland up in the northern part of Europe. That was a big, contiguous patch of Internet connectivity back in the days when the Internet wasn't that prevalent in Europe. Next slide, please.

There is a limitation in the DNS protocol. When you want to retrieve a list of the root name servers, you send a DNS query asking for precisely that. It was said in the early standards that the response should fit in a packet which is limited to 512 bytes, 512 characters. That's not a whole lot.

So a group of people – Bill Manning, Mark Kosters, Paul Vixie – sat together and thought, "Can we make this longer somehow?" They did so by just inventing new names for the hosts because they really know how the DNS protocol works, so they really know that inside these DNS packets, there is an algorithm that tries to pack the data together so it doesn't consume that much space in the packet. By carefully choosing the names, you could take advantage of that algorithm, and you could fit a few more in.

So the servers were renamed to the letters that we still have today. The plan was approved, and once that was accomplished, we had room for four more servers. Next slide, please.

You can see here how they were renamed. There had been a few changed to the lists, but this is actually more or less the list or root name servers we have, plus the four that were added this was completed. So you can see that nic.nordunet.net was named i.rootservers.net and it still carries that name. So this is still in place, this naming scheme. Next slide, please.

So four more servers were added, and Jon Postel, who was the then-head of the IANA, who is the central coordinator for the root service, used to set a criteria to select new root server operators. One of them was need, meaning the areas that weren't well-served by the existing ones. Another one was connectivity. The proposed host for a new server would have to show that they had very good network connectivity so it was easy to reach the server once it was placed there.

He also looked for a commitment to send and respond to traffic without any filtering of the content. Now, you have to remember that this was back in 1996/97. There was no secure DNS around, so he really wanted to have people he could trust. Also, where there was a community consensus, he would travel to

conferences and ask around, in Europe, “Who would be good to provide service in Europe?” He would go to Asia. He would go to various conferences in different places and around to form a picture of where the service would be well provided. Next slide, please.

This led to two servers being actually placed in different parts. In Europe, the RIPE NCC, which would [stay] there. The Network Coordination Center. Actually, the Regional Internet Registry for IP addresses in Europe was selected to run which is now K-Root. In Asia, the WIDE Project in Japan was chosen to run M-Root, and still does.

Two more were added: J-Root, which stayed at Network Solutions, which was to become VeriSign, and [N]-Root was transferred to ICANN when ICANN was founded because ICANN took over the role to run the IANA and it seemed like a good idea that they would also run a root name server because there are so tight connections between there. So that was put in place there. Next slide, please.

In 1998, unfortunately Jon Postel passed away way too early, and the root server operators who had up to that point had been working with Jon Postel as the central figure coordinating all the root server efforts realized that, “There’s no Jon Postel anymore.

There's no central coordination function. We need to do this ourselves."

And we did. We met as a group of individual organizations for the first time back in I think 1999. No, maybe late '98. This was a very secret meeting because we had no idea what was coming out of it. But it turned out that we all agreed on the most central, vital points of providing root service, the ones that we talked about before: the resiliency, using exactly the same information (the same root zone),and providing unfiltered information. All of that came out very quickly at that meeting, so we said, "Well, it seems that we are all on the same page. Let's continue to do this together."

There were a couple of more things we recognized, all the other root server operators. So among ourselves, we say, "Yes, you are operating F-Root. Yes, you are operating G-Root. Yes, I am operating I-root," and we all agreed on that.

We also decided to, if it would come to that one of the operators would no longer liked to provide the service, we would give very, very proper notice, well ahead of time, so that the other server operators and the world at large could adapt to that. Next slide, please.

So what does it look like today, and what are the features of the root server system? We are now 13 letters, and we've been that

for a very long time. Most of them, all except two, provide servers on IPv4 and IPv6. You can see here from the list who operates which server. We have VeriSign, University of Southern California, Cogent Communication, University of Maryland, NASA, the Internet System Consortium, the U.S. Department of Defense, the U.S. Army, Netnod in Sweden – we are operating the NORDUnet server – VeriSign, the RIPE NCC in Europe, ICANN, and the WIDE Project in Japan. Next slide, please.

Today, we are much more well-coordinated than when we were at that meeting after Jon Postel's death. There is a good number of servers spread all over the world. This is actually a dynamic map that you can have a look at it. If you go to the webpage www.root-servers.org, you will meet this map. It's a clickable map and you can zoom in. If you zoom in on Europe here, the 33 here will kind of expand and tell you exactly where these servers are located. Of course, that goes for all the other regions as well.

We are constantly working on improving the density of this map, each on our own merit. At Netnod, we go around. We talk to people. We want to deploy more servers. I'm quite sure that goes for the other providers as well. We're almost at 500 servers responding to your root queries on the Internet today. That's a massive amount of data capacity. Next slide, please.

So how is the root zone distributed to all these 500 servers? These are the 500 servers, spread all across the Internet. Each of these servers is operated by one of the twelve organizations, one of the twelve root server operators. It's located somewhere on the globe, and each operator feeds the root zone to these servers through their own distribution mechanism.

In the case of I-Root, we have distribution masters in Stockholm in Sweden that feed our servers in Wellington, New Zealand, in Singapore, in Tokyo, in Johannesburg, in Porto Alegre in Brazil, and in the U.S. – wherever we have them. And we feed our servers with the root zone.

Now, K-root in Amsterdam, they feed their servers with the same root zone, and we all copy it from the distribution masters that are provided by VeriSign. They in turn get the information they need from the IANA, which is operated by ICANN. But every change to the root zone that the IANA records or wants to have pushed into the system is authorized by the Department of Commerce and their National Telecommunications and Information Administration – yes! Got it.

And of course, what's the reason for making changes to the root zone? Well, that's because one of the top-level domains has made a change to their technical systems. If the Swedish top-level domain wants to make a change, they have to talk to the

IANA and say, “Oh, we want to buy a new service here for our TLD. Please make sure that that’s reflected in the root zone.” “Okay,” says IANA. “Is this according to policy?” It looks to the ccTLD policy for Sweden.” Yeah. Okay. Fine.” Then it sends off the request to VeriSign and NTIA for NTIA to approve and for VeriSign to implement it. It will then be distributed through the root server system.

This may look as a very complicated process, but I can tell you, once it ends up here… VeriSign, by the way, also adds the DNSSEC signatures needed in order to make sure that the clients can validate the content and make sure it’s correct.

Once it leaves here, it’s a matter of seconds before it’s out here. I would say typically less than two minutes before it’s out there and every client can get that new information. So if there’s any delay, it’s in the manual handling on this side. It’s not in the technical system here. Next slide, please.

We already talked about the diversity of organizational structure and the operational history as well. We all have diversity in hardware and software in use. We have diversity in operational models. We have diversity in funding models because this is a long-term undertaking. I’ve been working personally… The first time I laid my hand on the Swedish root server was in 1992, so I’ve been along for a long ride. So it’s a long undertaking, and

you have to ensure that you have the financial stability to provide this service for long-term.

We have common best practices for physical system security for how to provision. We have to make sure that we are able to respond not only to all the legitimate queries that come from all the users out there, but this system is occasionally attacked. We need to continue to work, even though we are under attack. You shouldn't notice. Have you noticed? I hope not, because then I'm not doing my job properly. If you notice that we are under attack, I am not doing my job.

So we have to over-provision. We have so much hardware out there that we're really, really working at being able to withstand or sustain every type of attack that we receive. I'm not promising that we can, but we are working really hard at trying.

We have also a very professional and trusted staff. As I said, I know most of the other guys in this business. We talk regularly. So if I have a phone call from someone who claims, "I'm root server B," "No, you're not. You don't sound like Wes. I'm sorry." So it's very tricky to infiltrate this community of root server operators because we know each other well and we've been doing this for a long time. Next slide, please.

Cooperation and coordination. We are much involved in various types of work. Many of us attend the Internet Engineering Task

Force, the standardization body for the protocols, and we're there to see what happens with the protocol and also give our input. It's happened in the past that people have come with new DNS ideas and we were forced to say, "I'm sorry. That will not work because that will put this type of load on the root servers, and we cannot withstand that." So they would have to rethink their ideas.

We are involved in the various regional operational environments, like the RIPE community, which is the European operators' forum. We are at NANOG, which is the North American Network Operators Group. DNS-OARC is an industry – I would call it a trade association, but it's an association for people who deal with DNS, not only root server operators. But there are lots of TLD people. There are researchers. There are all kinds of people that deal with the DNS, and they have lots of clever and good ideas.

Also, continuing here with APNIC and ARIN meetings, which are original operators meetings, and also AFNOG for Africa. We try to participate in as many meeting as we can, and we like to be approached. If you have any questions regarding root servers, come and talk to us, because that's the way we can reach you with our information.

There is also permanent infrastructure to respond to emergencies. We have set up communication channels. We all have our telephone numbers. We have communication bridges. We have chat systems. We have mail systems and what have you, and these are put on stable servers that are dedicated for root server operation only.

There's also coordination within established groups. I mentioned the Root Server System Advisory Committee here in ICANN. Again, IETF and OARC are places where we interact with the rest of the DNS community. Next slide, please. Also, hang onto this one a second. You will note here that none of these organizations actually deal with DNS policy. This is all DNS technology, and that's what we care about. Now, next slide please.

New requirements are put on the [DNS] system, and we follow that. We try to predict. We try to analyze. We try to feedback what we see into not only the operation, but also the standardization process and the processes that deal with policy because it can happen that the policy actually affects the technical operation, and we need to be in that loop and provide the feedback that gives the policy people an understanding of why their policy is impacting the technical operation and what negative effects it may have.

So internationalized domain names. For instance DNSSEC IP version 6 – how did that affect the root server system? Well, we look at it. We take help of DNS-OARC, which is actually the Operations Analysis and Research Center for DNS. We have resources looking at that, doing statistics, doing predictions, and so on.

We coordinate and work on evolving the robustness and responsiveness. The Anycast is our best and biggest tool here, but it has its own set of challenges, especially when it comes to Internet routing, making sure the packets go the right way. Next slide, please. Next slide, please. Okay.

There are a number of myths circulating out there regarding root servers. Let's try to have a look at a few of them. Root servers do not control where traffic goes. We have no control over your packets. The root servers is the thing you do before you do your transaction. If you're old enough to remember the number service in the telephone system, if you wanted to call someone and didn't know the number, you could dial a specific number and say, "What's Lars Liman's telephone number?" and the operator would look it up in a database and respond back to you with my telephone number. Then you would hang up. Then you would lift the receiver again and dial my number to call me.

That's exactly what the DNS is. You do the lookup before you do your traffic, and once you have done the lookup, the DNS system is no longer involved in your transaction, and your traffic goes through the routers of your service providers and other service providers until it reaches the web server, the voice server – whatever you want to do – the mail server. So root servers are not involved in controlling traffic.

Also, only a small fraction of all DNS queries are handled by the roots or even reach the root. Do you remember the caching? We talked about the caching DNS server that kept track of all the incoming responses to be able to make shortcuts. Well, these shortcuts make sure that most queries never hit the root. They only go direct to the server that needs to respond as far as the caching server knows. The more queries it receives, the more queries it needs to send, the more it learns. Eventually, when it's been running for a long time and has a lot of lines, it very rarely needs to talk to any name server at all. It depends on the clients using the machines, but if all your clients are trying to get to www.CNN.com, it will learn that very quickly, and it will respond from the cache. It doesn't have to talk to any server regarding that for a while, for a day or two.

Yes. The administration of the root zone is separate from service provision. The root zone is administrated at IANA, who gives instruction to VeriSign. The root server operators have no part in

that. We get the root zone from VeriSign. We copy them verbatim to our servers, and we provide the data that we are given. We cannot even modify it because if we do, secure DNS will immediately kick in and say, "Tell all the users that someone fiddled with this."

None of these servers are special. What you have is a system of 13 letters. Most of them map to a multiple instances, we call them, multiple copies, using Anycast. But all of these servers function identically. So ideally, you cannot see on the DNS response you're receiving whether you got it from F-Root or I-Root. It should give the identical response. If it doesn't, we don't do our work right. There is no difference.

The ordering of the letters has nothing to do with this. A-Root is not special in any way. B-Root is not special in any way. None of the others are. It's just a way to give them names. It doesn't signify that one is better or quicker or larger or anything.

Root server operators are not hobbyists. Definitely not. We have been working with this for a long time. For my own organization, we have two business legs, two business branches, that we stand on. One is providing DNS service. That is one of the big things we do. And we provide service, not only for root. We do provide it for other DNS clients as well, which means that it's very much the focus of the organization to provide good DNS

service. That goes for all my colleagues at all the other root server operators as well. The organizations are different. The focus is not. DNS service.

There are more than 13 servers. There are 13 letters, 13 IP addresses, but using Anycast technology, we now have opened hundreds of servers out there. So you no longer have to talk to one of the specific 13 machines. Forget that. That was 20 years ago. It's totally different now.

There's no single organization that controls the entire system. The root server operators are well-coordinated, but there is no organization on top of them that has the authority to control all 12. So we do coordinate and we get along very well and we provide a good service, we believe, but there's no one who can tell all 12 what to do. Try to create an organization that can tell the U.S. Army and Netnode what to do, that has the same authority over the two. It doesn't work.

Next slide, please. So if we now move focus from – let's see what time it is; yup – the root server system, the root server operators, and into the ICANN community, there is RSSAC, the Root Server System Advisory Committee. The role of RSSAC is to advise the ICANN community and specifically the ICANN Board on matters relating to the operation, administration, security, and integrity

of the Internet root server system. Now, that's the entire system, but it's only the root server system.

We are an advisory committee that has a very narrow scope. Compared to all the other advisory committees and supporting organizations, this is a very tiny bit of the system, and that's the only thing we care about. We don't care about the policies for new ccTLDs. We don't care about the policies for brand names. We don't care about the policies for – as long as it ends up in the root zone, it's our problem to make sure that everyone can get at that information and use it. But who gets what name? We don't care.

The only reason for us to get involved with a policy is if we see that it may have an impact on the service. If there's a new policy put in place that says, "Let's create a million new TLDs tomorrow," that will probably affect my service. I will probably say, "Would you mind going a bit slower so that we can make sure that we can ramp up our service in the same pace as you add new top-level domains?"

If the policy says, "Let's add this new technology to DNS [inaudible]." "Okay. Let me check how that will affect my service, because if it does it in a bad way, your customers and clients will suffer. So let's cooperate and see how we can make that happen."

The resiliency, the robustness of the system is our focus. We're not averse to making changes. But it should happen in a controlled way so that we maintain the stability of the network. Next slide, please.

So this is one organizational picture of ICANN. There are plenty, so next slide, please. We're located here. We're an advisory committee to the Board. RSSAC consists of representatives from the 12 organizations that operate root servers and liaisons to various other bodies.

We also have something called the RSSAC Caucus, which is a body of experts, which is actually a very open body. We very much welcome participation from all types of people – not only DNS server operators, but people who deal with DNS protocol, who deal with registries, who deal with general system administration and networking. There is room for a lot of expertise in that body, and we very much welcome that.

But we are an advisory committee to the Board, so we actually have a liaison to the Board, and we also have liaisons with the Security and Stability Advisory Committee, which has a much wider scope. They don't only focus on the root zone. They have a much wider Internet scope, so they're a much bigger advisory committee. But it's kind of a sibling committee for us, so we have a lot of common things to talk about. Next slide, please.

Yes. RSSAC is composed of, as I said, appointed representatives from the root server operators. Each has an alternate, or at least has the ability to have an alternate, in case the regular representative cannot make it to a meeting or cannot participate. The alternate can step in.

We have liaisons going to the root zone maintainers, to the IANA, the NTIA, and to VeriSign, because they are the ones who generate the root zone. We have liaisons to SSAC. We have a liaison to the Internet Architecture Board, which is the leading organization in the IETF circles for standardization. Our liaison to the Board is of course non-voting on the Board, and we also have a liaison to the ICANN NomCom, which is also non-voting.

RSSAC Caucus, yes. The members of the caucus are all volunteers. We really need them because the 12 of us that comprise RSSAC don't have the work capacity that we sometimes need to do research and produce documents. So that's where we involve the caucus. The caucus is formally appointed by RSSAC. Now, that is a very, very, very small threshold for joining the caucus. We haven't said no to anyone yet, so please join. Next slide, please.

The current two co-chairs of RSSAC are myself and Tripti Sinha, who's in the back. Would you please stand up? So that's Tripti.

The two of us are currently chairing the RSSAC. Next slide, please. I mentioned that. Yes, next slide again.

The purpose of the caucus. As I mentioned, it's a pool of experts who help us do research and produce documents. We need the expertise because the 12 of us don't have all the expertise we need in DNS and all the related areas. We need a critical mass of people. 12 of us are only 12 opinions, even if we can produce 14 if we set our minds to it. But we need a wider selection of people with more experience, simply a broader spectrum of opinions and experience.

It's also an important way for us to create transparency. It goes both ways. By having the caucus involved, we involve more people, which creates transparency. But we also are very transparent with who actually produces the work. So when the caucus is involved and produces a document, the document always lists the contributing members. So if you work in the caucus and you actually produce and help us create documents, your name will be on there because you helped us and you did good work.

But it also has the flip side, meaning that anyone who reads the document can see who helped create this document. We also have public statements of interest. Anyone who joins the caucus is required to produce a statement of interest where they

disclose their affiliations and they disclose why they want to participate in the caucus. Anyone who reads the document can go back to these published – they are on the web – statements of interest and see, "Oh, that Liman guy. Ah, he works for Netnod. He probably has some strange opinions about this. Have they colored this document so that when you read it you can assess and try to understand where people were coming from by looking at these statements of interest?"

The caucus is also a framework for getting the work done, so to speak. When we interact with caucus formally, we create something called a work specification. We make a request to the caucus, and we look for a work party leader who can drive the work in the caucus, who will then get a group of people helping him, and they will respond back to RSSAC with the outcome. And there's deadlines and all the ordinary mechanics trying to make the work happen there.

Currently, we have 67 people, and 42% of those are from non-root-server-related organizations. If you want to apply, just send an e-mail to RSSAC-membership@ICANN.org. It will be treated by the Membership Committee, which is currently Tripti Sinha, Paul Vixie, and Kaveh Ranjbar. Kaveh is the Chairman. He's in the back here. Go to talk to Kaveh or Tripti if you're interested in joining us. Next slide, please.

Publications. RSSAC was, in the first many years of its existence, not very well-known for making publications, but we're starting to pick up. We actually have our very first publication called RSSAC 000. When RSSAC we renovated, so to speak, we reshaped the organization two years ago. We started out by creating work procedures for ourselves to specify how should we work, what should our responsibilities be, and how should we interact with others. So that's already specified.

We have RSSAC 001, which is service expectations of root servers. When you talk about root severs, what do you really mean? What should you expect from root servers and root server operators? What is this service, really? And what should you not expect?

So RSSAC 001 is one part of this. This has been somewhat specified before in RFC documents (Request for Comments), which is the Internet Engineering Task Force way of publishing documents. There have been a couple generations of this, but the latest one in the RFC series is, by now, way over ten years old, and it's no longer relevant.

We also arrived at the conclusion that the IETF shouldn't really specify operational practices because that's not what the IETF is there for.

On the other hand, the IETF should have a say in the qualitative parts. What parts of the DNS protocol should you expect the DNS server to respond to, and how?

After some discussion, we ended up splitting this into two documents. So we now have the operational practices, which is specified in RSSAC 001, and there is an RFC about to be published by the Internet Architecture Board, which is part of the IETF service, which will look at the qualitative parts with DNS specification and so on.

The idea has been to release these in tandem, at the same time. Right now, the RFC document is being held up in the RFC editor's queue, waiting for an update. There's some negotiation between the involved parties on exactly what to update. But it's coming there. If you want all the gory details, come and talk to me afterwards.

We published RSSAC 002, which is an advisory on measurements of the root server system. The root server operators of course measure this system to make sure that it operates properly, and we do a lot of things. But we realized that we should probably agree on what to measure so that all of the 12 of us measure the same things so that we can compare the measurements between the server operators.

That's what this document is there for. It specifies a number of rather basic things, but at least we will be able to compare them and see trends. We don't really care about the exact numbers now, but we want to be able to see trends as this progresses. As new gTLDs are being deployed, as new technologies are being deployed, what does it do to our systems?

And we have a rather recent RSSAC 003, which is a report on the root zone time-to-lives. Each and every record in the DNS system has a time to live, which specifies the time that the cache is allowed to store the record before it needs to clear it out and go and fetch it again. This is the mechanism that makes sure that you don't have old data in your cache.

It turns out, if you start to look at these TDLs in very great detail, combined with the signature lifetimes in DNSSEC, you run into some interesting corners, where you could have problems validating DNSSEC because of timing issues where the TTL time and the signature validity time don't cooperate well. So we issued a recommendation to change the TTL values in the root zone, which has been picked up, which I think is possibly about to be implemented. It was at least well-received.

We have made a few statements as well. We have made a statement on the ICG proposal on the CCWG Work Stream 1 report. Also on the IAB liaison to RSSAC, we have a mutual

exchange of a statement about how the IAB and RSSAC are supposed to interact.

When that TTL problem was discovered, we so to say made a quick fix. We made a statement to change the signature validity period in the root zone first because that's actually a smaller thing than changing the TTL. Then we launched a research party to figure out what the proper fix would be, and that came out in RSSAC 003. So these are all fairly recent documents. Next slide, please.

Current work status. Let me see what it says here. Yes. We have chartered a work party, so we're talking to the caucus and asked them to create a work party to create the history and technical analysis of the naming scheme used for individual root servers.

So this is to document the technical history of the names designed and consider changes to the current naming scheme. I mentioned that we've had this b.rootserver.net scheme in place for over 20 years now. Can we make changes to that in order to make improvements to the system? Or can we today safely ignore that 512-byte limitation that was put in the specification 30 years ago? And please also do some risk analysis regarding that, and eventually at the end, make a recommendation to root server operators and root server management partners on

whether changes should be made, and if so, what the changes should be in order to improve the service.

This work party meets weekly and expects to finish its work in mid-November, reporting back to RSSAC. Next slide, please.

I think that we are nearing the end, aren't we? That's the very last one. Okay. That's the end of the presentation then. Now I would be very happy to receive questions or comments. I see in the back several representatives from other root server operators who I'm sure will step in if questions go that way.

DAVID MORRISON: Hi, Lars. I'm David from New Zealand, and I'm losing my voice. Question around when the cache communicates to the root. What determines which server responds to the cache when there's hundreds and hundreds of servers out there?

STEVE CONTE: Right. Good question. So we have the picture where you have the end user. You have the caching DNS server, and then you have the root server at the top. I showed only root server in that picture.

So how does that caching DNS server now which root server to talk to? When it starts, the cache is empty. There is no information there. So how does it know about the root servers at all? Well, each and every one of these caching DNS servers

actually has a configuration file or corresponding information that lists the root servers, all 13 of them, because there are 13 IPv4 addresses and I believe 11 IPv6 addresses, and they have to be there in the configuration file.

Okay, so when it starts up, it knows about the roots of operators, but it still doesn't know which one to talk to. This is actually, to be honest, software dependent; depending on the software that you run in your caching name server.

But one method that is being deployed – and I think by more than one – is that it will choose one at random first, and it will send the query to that random root server.

Now, for each one of these 13 IPv4 addresses, or for most of them, there are a number of servers responding from that IP address. It's actually up to the routing system, not to the roots of operators, but the people who operate the network, to carry that packet to the nearest one. Now, mind you, "nearest" is a very strange definition on the Internet, but there is a notion of "nearest." All the routers out there know what the nearest one is. A very strange mathematical view of "nearest." So it will send it to the nearest IP address for that IP addresses that its caching name server chose.

Next time it needs to talk to a root server, it will choose another one, and then another one, and then another one. Eventually it

will have [walked] through all 13. For each of them, it will record the time it takes to get a response. When it's walked through all the 13, it will know which one is the quickest one, which doesn't mean the nearest one. But that doesn't matter because quick is of essence. Distance is not. And it will hang on to that quickest one.

But sometimes things change on the network, so eventually from time to time, it will do the 13 walk again to make sure that it chooses the quickest one.

UNIDENTIFIED MALE: Hello. I'm [inaudible] from India. I want to know who appoints root server operators. There are 13 technical entities. Is there any plan to increase as the Internet use is increasing?

LARS LIMAN: If you're taking them one by one, who appoints a root server operator? Jon Postel does. Unfortunately, he's dead. So there is currently no process for appointing root server operators. At least I'm not aware of one. You can invent one in your backyard, but there's no one who is widely known and universally accepted.

There is pressure to increase the number of root server operators, and RSSAC is aware of that. So I believe that work has

to be undertaken to see if we can find a way to expand the numbers. But we also need to make sure that we expand the numbers for the right reasons.

For me, as a root server operator, the reasons to expand the numbers would be if there are technical reasons, if there are people who don't receive the responses as they should, if there are people who see intermittent service, where there are interruptions, where it doesn't work, where there's a delay in conveying new data. That would be reasons for me to look to establish new root server operators if we could find someone who meets the criteria.

We first have to define the criteria. What are the criteria for becoming a root server operator? So there is a ton of work to do in order to create that process. It's nothing you can do off the back of your hand. It's something that will have to be carried out very carefully and very slowly.

UNIDENTIFIED MALE: Thank you. Great presentation. I appreciate the clarification of all this. What risk is there at that caching server that that does not become the authoritative information? It's related. I know it's not your responsibility, but what responsibility is there that the caching server doesn't get corrupt?

LARS LIMAN: Right. Is there any chance you can put that slide back up? Because I think we have several questions. The one with the end user and the caching name server.

You are quite right. There's a risk that that server provides incorrect information back to the end user. That's definitely the case, but mind you, the end user has access to all the DNSSEC information, so the end user can actually perform the validation. It requires that the end user is, so to speak, knowledgeable and has the right software tools.

But the protocol allows for that. There's nothing in the way here that prevents the end user from doing that calculation and making sure that he gets the correct data. That said, that's probably not the general case.

And, yes, then you have a problem with the caching name server giving incorrect information. But as an end user, you often – not always – have a choice. So if you start to suspect that your ISP, which is typically the provider of this service, gives you incorrect information, at least on the working market you can choose another ISP who hopefully performs better.

To be honest, you could run this service yourself. You could run your own caching name server. I do on my laptop. It's not that

hard to set it up, but it probably requires you taking a course in DNS administration.

More questions? Ed, we have a question up here.

UNIDENTIFIED MALE: Thank you. [inaudible] with GAC. I was wondering if you would be so kind to also share a bit about how the operators are funded.

LARS LIMAN: I don't know. I can talk for Netnod. The others would have to speak for themselves. I guess we have a few here, and I would gladly see you speak for yourselves.

In the case of Netnod, Netnod has two businesses branches. One is to operate exchange point for Internet Service Providers. The other one is to provide DNS service.

On the DNS service side, we provide the root service, of course, but we also provide DNS service for top-level domains and other high profile domains. That is a service we charge for. We don't charge for the root service because there's no on too charge.

But in the TLD case, there's a TLD registry we can charge. We charge for the service, and the money we charge is used to provide that TLD service and the root service.

We're in the fortunate position that TLD operators are often positive to this because the TLD really depends on the root to work, because if you take out the root, the TLD doesn't work. So we have no conflict here in interest.

So for us it works out. We provide our service, and we also have the exchange point where we have a revenue which we can use to cross-subsidize the root service if we have to. But it's mainly dependent on the DNS service we sell.

Other comments from others, please.

UNIDENTIFIED FEMALE: In my case, which is D-Root (University of Maryland), we're self-funded. The university picks up the cost.

TERRY MANDERSON: I'm Terry Manderson from ICANN. L-Root is funded by ICANN, and obviously, ICANN is funded by all of you.

UNIDENTIFIED MALE: [inaudible] K-Root. We are a technical membership organization, RIPE NCC. So basically almost all ISPs in Europe, the Middle East, and Central Asia are members of us. K-Root is also funded by membership fees.

LARS LIMAN: Thank you. Any more questions? Are there any questions in the Jabber room?

UNIDENTIFIED MALE: No. No questions in the Jabber room. Thank you very, very much.

LARS LIMAN: Thank you. I should also thank the people who have produced these slides. These slides are not my own work. I know that Duane Wessels has put a lot of effort in this. There are a couple of other guys who helped produce these slides. I should also mention Steve Sheng, our staff support, who helps us a lot with all these practical arrangements. Thank you so much.

Thanks to all of you for coming here. And again, come and talk to me.

UNIDENTIFIED MALE: Now I would like to ask the audience to just to please take a moment. We have a little survey. We're trying to figure out what kind of sessions we need to do next. Your opinion and perspective will really help a lot. It's a very short survey, so if you could take a minute, please go ahead and fill it out for us. Thank you.

[END OF TRANSCRIPTION]