MARRAKECH – Tech Day
Monday, March 07, 2016 – 10:00 to 16:45 WET
ICANN55 | Marrakech, Morocco

UNKNOWN SPEAKER: …now we're defining words.

This is ICANN 55, Tech Day meeting in Oliveraie. On March 7th from 10 to 16:45 WET.

EBHERHARD LISSE: Good morning everybody. Can we settle down please? Excuse me, can we start or do we need to finish our deliberations first? All right, good morning. This is Tech Day at ICANN 55. I'm Ebherhard Lisse. I'm reasonably well know so that I don't have to introduce myself again.

It's, I think, the 27th Tech Day, and we have been struggling slightly with two cancellations, one only yesterday, but fortunately Christian Hesselman was willing to step in on very short notice so I could spare you and didn't have to give you my presentation about the Zika Virus, which, as you know, I have done in my professional capacity for my colleagues in [inaudible].

I have taken the science out so that it is good for lay people, but I thought it maybe a topic on point would be better. As usual, a quick rundown on what we're going to do. Ondrej Sury is going to give us an update on developments with their name server. Then Marvin Woo is going to talk a little bit about IDN email, the Chinese Core Mail seems to be driving the, spearheading this in many ways.

Francisco Arias, is he here? I haven't seen him yet, is going to speak about EBERO, well that's the facility that ICANN has implemented with support of two or three back ends. One of them is nominate where if a gTLD fails, they have a method to sort of sort that out.

They were able to test this on the live TLD, dot Doosan, was a gTLD by a company that was basically not used and still relinquished it afterwards, so they can actually use, see how hard this works in live, and they learned some lessons which they can talk to.

Then we, Attila Ozgit, is he here? He's there. So in case Francisco doesn't come on time, we may have to move you a little bit forward. Will speak a little bit about the DDos attack that happened last year on dot TR.

I want to mention right in the beginning that this is a friendly audience here. It's obviously difficult for somebody to come and

explain how they were attacked, so the idea is that we all learned what happened, that we all learn what they did to mitigate it, and that we can all learn from how to prevent it from happening.

And if that doesn't work, that we can learn what works, what we can do to prevent this. Then Christian Hesselman stepped in on short notice with the [inaudible] TLD presentation. Jeff Osborn from IFC will talk about the F-root deployment strategy.

And yeah, Jaap Akherhuis asked me for a short, a small spot on quick and dirty DNSSEC monitoring. He asked me a few months ago when dot BW had a little issue with their DNSSEC, in order to contact, and I was unable to, I even spoke to one of the people from dot BW yesterday, but they couldn't help me. And then he mentioned that he has a few [share?] clips, and since it's easy to implement situations and solutions.

So I asked him to show what he's doing. Let me move forward here myself, preferably. Somebody is interfering with my success here. So in the afternoon, Don Hallander, we had to rearrange the lunch time a little bit because Don Hallander was having a full, I see him in the back, but he originally was going to have a full day meeting in the morning, or at least he couldn't do his presentation in the morning.

Maybe his session starts a little bit later, but in any case, we were a little bit tight by availability of speakers, so we had to do the lunchtime a little bit out of normal timing.

Ricardo Scmidt from University of… Is he here? There you are. How are you? Will talk about any casting, and there also we had a little bit of an issue because we have an any cast roundtable. We wanted to keep his presentation separate, but not equal, but separate, but connected to it. So that he could also then be stood here when the any cast round table start to share his insight on what's being discussed.

And then Jacques Latour will do the usual closing remarks. So, first we have, and I must say, our lunch time will be self-catering, unfortunately. We haven't really figured out a budget for this yet, but I think will do. And Andrew Sullivan will start with the KNOT version two.

We will present from this laptop. We can advance the slides, or you can advance the slides yourself. Or you can ask for the next slide and it will be advanced.

ONDREJ SURY: Does it work? Yeah. Hello. My name is Ondrej Sury, and I'm from CZ NIC, and this presentation is in fact about our new project, KNOT resolver. The, well, the previous one was not DNS,

which was the operative DNS. And this is the records of DNS server.

So what is the KNOT DNS resolver? It's not only resolver, we view it as a platform for building recursive DNS services because it's so much flexible. It's open source, it's all work we do at CZ NIC. And it has full DNSSEC support, including the [inaudible] curves, including the RFC 5011, which is automatic trust anchor management.

And including the negative stress anchors, which is quite new RFC. Next slide please.

This is not next slide.

Thank you. So, it's written in C and Lua, which is very fast and better language. And it's scripted because the configuration file is in fact the Lua script. So you can program things inside, and it allows the Daemon to, or the operator to reflect the current situation much more flexibly than just what the vendor releases.

So you can, you see a new threat, and you can just snippet of code, it will prevent the attack. It has a very simple core, which is more extensive in the [inaudible] version in C, Lua, and Go language. We have support for happy eyeballs, IPv6, that gives 20 milliseconds SR for IPv6. And it has no internal threading. It's just one process, and it scales by replication for up new process.

And with the recent kernel feature, which is called [SR U Sport?], it will just bind to the same port and answer, the kernel will take care of distribution of the queries between the Daemons. Next slide please.

Still no next slide.

Oh, thank you. So who is this for? Well, it's for everybody. It's suited for large DNS farms, and I will tell you why in a moment. It's also suited for small recursors in private network, and for personal resolvers on your laptops, and for all geeks, tinkerers, you, and people who like to play with DNS things. Next slide please.

So why large recursive DNS farms? Because it scales, and really fast script in Lua allows you to change resolution on the fly to reflect your situation. Is this a flexible shared cache back ends? One of these is local using LMBD, this is lightening memory database which is probably the fastest open source key value storage there is.

And we have two networked cache back ends, MEMCAHED and REDIS is also very fast. So network nodes can share the same cache for the DNS resolving. And the new instances, if you need, well more power than you just fire up a new instance, and it will just pick up the shared cache data, and the process will start with cache, so the resolution DNS resolution will be fast.

It has great statistics, metrics, and plotting because of the Graphite back end that can feed the data into, for example, Influx DB and plotted the data with Grafana, and it has negative trust anchors. It's cluster aware. There is ETCD module for shared self-configuration, and it can [inaudible] to ETCD daemon and pull up the configuration that is shared inside the cluster.

It has support for views and ACLs, and it was a prefetching. So the cache stays hot before the [inaudible] expire. Next slide please.

So this is just plotting in Grafana so you see how nice graphs can be made from data from the resolver. Next slide please.

So why small recursors in private networks? I mean, the home networks, the company networks. It could be well, a small to medium networks, even router I would say.

It has QNAME minimization, which is the recent sender in IETF for DNS privacy. So who doesn't know what the QNAME minimization is? So who doesn't know? Okay, so everybody knows. That's great, that's great. Okay so I will not explain it. It has a good DNSSEC and RFC 5011 key management. It has low memory consumption because the cache can be paged out to the [inaudible].

And it has query policy based resolution, so you have flexible filters based on pattern, suffix, responsible zone. And you can do all of this stuff like deny, drop forward, [inaudible] the query to TCP with the CC flag. And it also has DNS 64 support to complement NAT 64 for IPv6 only networks. Next slide please.

So why for personal resolvers? It has a very simply config-less operation, you will just give it a writeable file to DNS trust anchor, or resolver can download a key from the IANA using the certifications. And you are good to go. That's all you need to do, to run KNOT resolver.

It has persistent caching because LMDB, so it survives reloads, restarts, reboots, of the daemon. And it has a tiny web module for monitoring your queries. There is a live demo at the URL, or I will just show you on the next slide how it looks like.

And for future work will improve even more privacy with DNS over TLS, and DNS over HTTP sort of to prevent all of those, also a Wi-Fi on the hotels that will intercept your DNS queries and break it. So this is the output of the tiny web. It will just show how fast the resolution of names are, and geo IP map of the servers queried.

And even below there is a list of queries. It is very common [inaudible] if you are compiling something, I just look at this and see how my resolutions are. Next slide please.

So for Geek thinkers, why? Because KRSED, thisi s deamon, is scriptable without binding on port 53, so you can run all sorts of [inaudible] queries using [inaudible] API. And there are even example scripts. One is KRSED host, we sort of like the host utility.

And the second one is the KRESD query, which you can exploit full API in the LUA to do all sorts of things with the packets, with the queries, with the answers. The examples here will just bring the query name of the response, and our code of the response of the query you issue to that.

The full API specification is in the documentation, which is on the next slide, I think.

Yeah. The resolver comes with really good documentation, the URL is there. And we are in the beta phase, almost coming to release candidate on version one of 1.0, and we are doing testing to be really sure that everything resolves as it should because the DNS is really broken. And there are a lot of broken service out there, which will just break the resolution, and you need to do sort of lot of work around to make it work.

So I would like to ask you to give it a try. We have a new website for that, where you will find all of the resources for KNOT resolver. We have the Debian and Ubuntu packages of [inaudible] to come. And there are source code, because it's

open source as I already said. And even if you are really, really lazy, you can use Docker to issue just one comment and it will come up and run.

So through a normal DNS and really weird DNS stuff on it, and report back to us if you find any bugs, and even if you are successful and you are very happy with the program, we would like to hear from you. And I think that's all. So if you can show the GIF.

Meanwhile, if you have any questions. So that's on our very good choice of name of the product.

So any questions? Even you can KNOT.

ELBERHARD LISSE: I can tie a knot, I can tell you that.

ONDREJ SURY: Okay, Jeff.

ELBERHARD LISSE: For the remote participants, can you give the microphone to the…

JEFF HOUSTON: Jeff Houston, APNIC. You said you had RFC 5011 support, one of our concerns in looking at that, was that when we roll the root key, and we're relying on RFC 5011, there is this issue of a 30 day hold down period. But if I turn on my resolver, in the middle of that period, I won't see the new key necessarily for 30 days.

Have you allowed any manual override to learn that new key faster than the 5011 mandated 30 days?

ONDREJ SURY: I think it's quite easy, if you just delete the content of the file. It will just re-download the new key from the IANA website using the certificates.

JEFF HOUSTON: So it should just pull it down without necessarily relying on a 5011 roll. Okay. Thank you.

ONDREJ SURY: Warren?

WARREN: Warren [inaudible]. Sorry I arrived kind of late. Did you say it will automatically pull down the key from the IANA website?

ONDREJ SURY: Yup.

WARREN: That's awesome. [LAUGHTER AND APPLAUSE]

ONDREJ SURY: A little exercise in the morning.

DAN YORK: Dan York, Internet Society. So what's next for KNOT? What are you doing next to it?

ONDREJ SURY: [Inaudible] diminish. [LAUGHTER] I'm not sure what you're asking for.

DAN YORK: What is your roadmap?

ONDREJ SURY: Our roadmap is to finalize the current release, and do more testing. And I think that the next stuff on the roadmap, the bigger things is IT cookies, DNS cookies. And the DNS over TLS will be [inaudible] also this year. But we would like to wrap up

this production really. To really release the production of this version shortly, basically after we finish more forward testing.

And switch my microphones.

EBERHARD LISEE: Any more questions? Anything from the remote? All right. Thank you very much.

So fortunately we're running a little bit ahead of time, so Marvin Woo can make his way to the seat. Why don't you come and sit, or you can stand if you want, whatever you prefer. You can sit or you can stand and use the microphone, whatever you prefer.

You can advance your slides yourself.

MARVIN WOO: Hi everyone. I'm Marvin Woo from Core Mail China. This is the first time I show my [inaudible] in the Tech Day. Thank you.

Okay. My presentation includes four parts. Who is Core Mail? What's IDN Email? EAI commercialization progress, and some challenges and plans on EAI.

Okay. Core Mail is a company to email only, doing 17 years. We are only to email. We were founded in 1999, which is the same year as the IDN. So we [serve] more than 700 million users,

because Chinese has too many people. So we can get a lot of users.

And this picture is our management team. Okay. So my [inaudible] is who is me? In the right picture, [inaudible] is me.

Okay. Our team is funny and passionate. Okay. What is IDN email? IDN email is email that contains international characters, such as Chinese and Japanese and so on. Now, our email address only contains English number, not likely now. EAI address can work like this picture.

And we use some protocol and standards, include UTF-8 is our encoding. And some standards are RFC 6531 and 6532, 6855 and 6856. And that is a problem, our email system may need to reconstruction, because before we support EAI, our system only can accept ASCII encoding, ASCII code.

So our inner system need reconstruction again to change it to support Unicode. And we solve our custom tools and to double email address. And such as this. Before I sent email, the custom made choice which address he used, but there are the same email address.

Just like me. My Chinese name is [foreign language] and my English name is Marvin Woo. So they're all the same person, and

our email solution is like this. Chinese or any international user name, and alias name is the send account.

And so we solve double email address like this, for example, our main email address support UTF-8. And we give the same address alias, alias name like English name. So, they are working as the same account.

And when receive email, the system will do a judgment. If the receive can support EAI or not. If yes, we send Unicode email to the receiver. If not, we give him ASCII code. Okay.

And the sender is also maybe [inaudible] do a judgment if the, accept and can support Unicode or not. So it's less send [inaudible]. Okay. So things to 2012, Core Mail has the EAI solution, that we found no applications support EAI and no email operation can support. So we can send the email with ourselves, but it's fairly trouble.

Now, Google, Microsoft can support it. But in 2012, not only Core Mail could support it. So we have too many challenges. Now I can experience something for us. Okay.

And the most challenging is no other email operation can accept Unicode. And some users, [inaudible], interesting, so interested in the EAI, but also ask me, but we have no place to register EAI account, because Core Mail is email system provider. We are not

an email operation, operate. So we release EAI register platform with China's [inaudible].

And we found our users, if he has EAI account, but no client can support EAI because our client, like Outlook, like [inaudible] mail, and so on, only can accept ASCII code. So Core Mail, we release some client name, flash mail, and we support some mobile client, name is Core Mail APP and Core Mail link.

So we have fuller solution like Windows client and mobile client. And if some, an enterprise customer need EAI account, so we upgrade our [sass] platform to several, to give our customers. This interface is planned from like people can accept some EAI account. If you have… You can try. It can be free to register.

Now, there will be more than 100,000 users in this platform that are free. And there also is a [inaudible] EAI account. The name now can be [foreign language]. Is real Chinese name, but also we can support other language.

So these are client's name. Core Mail's APP, Core Mail's flash mail, Core Mail's Lunkr and we can support IOS, Android, and Windows. The right picture is our interface. Last year, we have Thailand and India to make an EAI platform, and to support the Thai language and Hindi language, is India and Thailand. Now they can work now.

Okay. And in Chinese, we have some commercialization progress, because [inaudible] may be to be used could make benefit, it has okay. So in last year, we make some benefit for EAI. The two customs [inaudible] is registry that build [inaudible] EAI solution, use commerce system, and they pay for me. So it can make benefit and commercialization.

And our commerce planned for [sass] users also can use two account. One is EAI account and the other the English name. Okay.

So, the challenges also I said, no few APPs can register because now, email also accept to send and receive email, also email ID as registered ID to use. Example, this ICANN meeting our register ID is our email address, but you may found also they only can accept English name.

I try my Chinese user name, but the system can't remember, can't accept. So, this is a fairly big challenge because email address also use to, as register ID such as online games, mobile games also use to register ID. But all of these systems can't accept EAI, can's accept Unicode.

And so [inaudible] the challenge. Okay, thank you.

In the future, we think we also can [inaudible] EAI email in China. And the other non-English countries that are in corporation with

some [inaudible]. If some people have interest where we can connect me. And I think sometimes a close alias will be okay, such as any government also use email in closed areas. Maybe we can do these areas.

Okay. So thank you.

EBERHARD LISSE: Thank you very much. We can all give him a hand. [APPLAUSE]

I have a question from the chair. I don't really understand, if I send an email to any EAI address, how does it find a way to your system? If I, for example, send it to Russia, how is it going to work? I really don't understand if my mail client understands it, how does it do the resolution to find the MX record and how does that work?

How does this interface with the DNS that I haven't figured it out yet.

MARVIN WOO: Interface and address is the two things. And the system will be judgment, whether the address can support or not support. If the address can support EAI, we can send ASCII code to the receiver.

EBERHARD LISEE: Now, if I'm trying to send an email address, email to an EAI address, in a country, let's say, to an Arab speaking address. How does the DNS connect this, I don't really understand how the mail connects to the DNS, because you basically need to go and find your way to your server.

I mean, I get emails from you, you send it from your Chinese address, and it comes up in my ASCII address, I can see it here on the screen. So from the sending side it works. I haven't really figured out conceptually how it works the other way around, if I was sending from Thunderbird or something using a Chinese address, how it would find a way to your system.

MARVIN WOO: Your meaning, DNS how to classify. IDN also can work, still now it can be work now. IDN, the domain name DNS works for domain name. IDN now is can be work. DNS no, it's okay. Okay. So, the email address is work before under the IDN. IDN is not ready.

EBERHARD LISEE: Any questions on the floor? Any questions from the remote side?

UNKNOWN SPEAKER: Yes, we have one question from Calvin Brown. I'm going to read the question. "As I will never be able to understand with these emails, I'd be interested in ways to responsibly block them. Are you working on incorporating this spot in your offering?" End of question.

MARVIN WOO: Hello Mark. Can you help me? My, Microsoft, yeah, I know.

MARK: So this is Mark from Microsoft. I really don't understand the question, just because email is sent to an internationalized address doesn't mean that you can't read the content. The content could be in any language and in any script. It's just that now the address owner has an address in their own local script.

So the first question is, I don't think a priority you want to block all EAI emails anymore than you would want to block all dot XYZ TLDs. You remember there was a few months ago, someone on the Internet said, "Oh, you should block all XYZs." That was terrible.

EBERHARD LISEE: That's not the question. The question is, I think, he's in [inaudible] on the sick bed, he's sick. Otherwise he would come,

he usually comes here. How can you block spam? But them in a spam filter? How is this working with Chinese language? I have no clue. [CROSSTALK]

UNKNOWN SPEAKER: …also if there is a question from another way, from what I understand is, when I use my email client, I have all kinds of rules to sort out my emails, and I can see from my boss goes to priority, my wife goes to priority, from XYZ, maybe later.

If there is an email address that comes with from that has Chinese encoding characters, and I'm in the US with my basic version of Outlook, how does it work?

EBERHARD LISEE: Or to make the point, if I get email from MWU at Core Mail dot com, I know it's him, it goes. If I get from his same address in Chinese, I have no idea. How do I…? Now if I wanted to filter him, not necessarily as spam to block, but just to filter him into a folder, how would I do that?

MARK: So, assuming your email client, you know, supports UTF-8, then parsing these strings is no different from any other address. And we are ensuring that these features are put into Office 365, the

anti-malware, anti-spam front is called EOP. There will actually be a discussion about EOP and EAI at, oh. Lars is gone. In Cologne, in April. I forget the name of the conference, we'll be talking about that.

And then within the Outlook client, for instance, similar behavior. You know, you have to recognize that the EAI address, or the IDN version of the address are actually the same address, and so you have to be able to do those conversions on the fly when you're doing searching and sorting. That's a feature is being scoped. We don't have that yet.

I'm not sure if I'm answering the question. From our perspective, as long as you understand that these are Unicode strings, you just treat them all the same. And so there is searching, there is sorting, and then spam filtering.

EBERHARD LISEE: Can you answer? Or have you any comment?

MARVIN WOO: Maybe Duncan, say anything?

EBERHARD LISEE: Very diplomatic. Don Hallander.

DON HALLANDER: Yes, it's Don Hallander here from the UASG. And to answer Calvin's question, I think we're befuddled by the question because it's just another email address. It just happens to be in Unicode, or it happens to be in a script that you may or may not understand.

Already today, and for the past forever, we have been seeing email address displays coming up in characters that you may not understand, even though the address is an ASCII address. And Eberhard, I'm not sure if you got an answer to your first question in terms of how do you find the EMX record, but you just do a Punycode conversion.

So you take the Unicode address, which is in terms of the domain name, convert that to Punycode, send that over the wire, and that finds the DNS record and points you to the right address for your MX record, and mail goes through.

So the issues with EAI are two-fold. The hard bit is, how do you deal with a user name or a mailbox name, or the local part name, or the name to the left of the at sign in left to right scripts? How do you deal with that when it's not in ASCII? So that's the hard part.

The easier part, is how do you display IDNs properly in your email client, and that's just the conversion. That's pretty straightforward. But, very few people are doing that, and that's

what the universal acceptance steering group is pursing is to encourage mail operators to [inaudible] code. Does that answer your first question? And Calvin's question?

EBERHARD LISEE: I don't know whether that helps Calvin's question, it helps my understanding of Calvin's question.

MARVIN WOO: Another thing is maybe, yeah, it's in [new things]. Maybe the email address is yours, not for us because all, a [inaudible] can't understand English. I think maybe [lease rooms?], no need to know need use to EAI, because we all know English. We all can pronounce English [inaudible], but a lot of people can't.

Remember, it's English name, like my father. Like my father, he's Chinese old farmer. He can't understand English, but I give him a Chinese account. He remembers now, soon but I do email more than 10 years, he didn't never use my product. Here I give him an EAI account. So I think EAI not use, not for us, for some people, less knowledge to use email, to touch Internet.

EBERHARD LISEE: I think the same point your colleague who said, it's absolutely clear. I mean, I can send you an email no matter what, but your

own father, or mother-in-law, or whatever from the rural areas. My mother-in-law doesn't understand English, fortunately she would use, she can realize my name because it's in the same script.

But I fully understand that for someone who doesn't speak English, it's very difficult to communicate by email with their own sons who live in the town, in a big city, because they just… You can type your message in Chinese, that will work.

If necessary, you send it as an attachment, but you can't, you don't know how to tap the address. I cannot tap a Chinese address, if that was the only choice. So that's… I understand, and it might not only be for, it's not just for China. It's the same in any other country, whether it's an especially large population that doesn't understand the language, but they have access to computers, and a part from use of the computers, because they just don't know how to address email to people that they want to communicate with.

So that it makes all sort of sense. Thank you very much. [APPLAUSE]

UNKNOWN SPEAKER: There was a comment, I have a comment from Calvin who said his question was more about accountant, it's in the [inaudible]

that he can't understand and he simply don't want it. He would like to block it at SMTP transaction level. That was a comment, not a question.

EBERHARD LISEE: We'll defer the question for next time we see him. Okay, Francisco Arias will be next, talk about EBERO. And why doesn't Simon come on the panel as well in case there are questions or contributions that you can…?

Francisco, it's on the presenting laptop already so we can…

FRANCISCO ARIAS: Hello everyone. This is Francisco Arias from ICANN staff, and to my right I have Simon McCalla from [inaudible]. This is a short presentation on, excuse me, an exercise that we jointly did a few weeks ago. This is the emergency back end registry operator. I'm going to explain what this is about.

This is the agenda for the presentation, and let's go into the matter. So what's the EBERO program? This is a program within ICANN. This is related to the new TLD program in ICANN. So all of these close to 1,000 new TLDs that have been added to the root have clause in their contracts that allows ICANN to take over the operation of the TLD if they reach certain thresholds of availability in each of the five critical functions.

I'm going to explain later what these are. So intention of the program is to protect the registrants within the new gTLD program from registry operator failures. So if DNS goes down, for example, for four hours, then we can take over the operation and restore that service as soon as possible.

And the, we have three companies that we have contracted for this service within ICANN. And they are, the main qualifier for them is that they have use of experience of operating registry services. That's the main characteristic that they have. And we have, of course, a contract with them, so we can have isolates in terms of the things that we expect from them.

These are the five critical functions that I was referring to. They are laid out in the registry agreement, and they are also what we request from the EBERO providers is DNS resolution, proper DNSSEC operation of the TLD. The operation of shared registration system meaning EPP. And we also require them to provide we call in ICANN the registration data directory services, that's WHOIS, the port 43, and web based WHOIS.

And finally, we also request that they either provide what we call data escrow, that's on a daily basis they provide backup of the registration data to a third party that we have contracted to hold this data, should there be a failure in the root itself.

So these are three EBEROs we have, the CNNIC, that you probably know, CORE in Europe and Nominet.

Now I'm going to talk shortly about how the EBERO event works. So what is an EBERO event? That's when a TLD fails, or when the thresholds are met. So we have a monitoring system that we operate. This monitoring system is constantly checking the DNS and WHOIS at the moment. Eventually we're going to add, we plan to add EPP monitoring.

We monitor all of the TLDs, and we basically looking for a response time. If the respond at a certain threshold, there is a final contract, then we consider the service down. And so the, we have a team within ICANN that is operating 7 by 24, and they are looking at these alerts. And if there is an issue, we first try to resolve the issue within the registry operator if that's not possible within a certain time, then we start following those procedures and eventually, if necessary, we declare an EBERO event, and have one of our EBERO providers to take over operation of the TLD.

So far, we have had zero real EBERO events in the program. In other words, there have been no need to migrate the operation of a TLD to one of our EBERO providers. This is what would happen in an EBERO situation, I guess shouldn't be any surprise to all of you.

We in ICANN have copies, daily copies of the TLD zone files for all of the new TLDs. So as soon as we declare an event, we will make this zone file available to the EBERO provider, wherever the EBERO provider turns out to be. And we ask them to immediately sign that zone that we have available, and put it up in the DNS infrastructure.

Once that's done, then we would request a change from IANA to re-delegate the TLD, and once that is completed, then we will have the service up in the DNS living inside caching effects, and they were talking about a transitioning with the back, or register operator. It's not cooperating then there will also be DNSSEC issues temporarily.

On the side of the other services, the SRS registration service and the WHOIS, what we do is all the new TLDs have data escrow agent that they have identified, also as part of that, an agreement that they have with their data escrow agent, and they identify ICANN as a beneficiary of the data. And as in that contract, they have to put certain clauses, basically if they reach the emergency threshold, then we would have access to the data.

So in the case of an EBERO event, we will call the data escrow agent and request the release of the data. The data escrow agents also have a SLA, they have to turn on the data, turn back

data within 24 hours. We have the data, then we made that available to either all of these properly encrypted and authenticated, the communications between us, the data escrow agents, and the EBERO providers.

And then on their side, on the side of the EBERO provider, they have to import the data, transport the data with the zone file. We have a procedure for those validations, basically whatever is the newest data, be it from the data escrow, the deposit or the zone file, is the one that is considered authoritative if there were any discrepancies between the two sources of data. And that's the one that will be taken into consideration.

So with that then the EBERO provider proceeds to enable the RBS and EPP service under escrow. Once the situation is considered stable, then we have… In a normal situation, we would not allow dates, sorry not dates, registration of new domain names in a TLD that is under EBERO operation, but we have operation there to allow ICANN to request emergency updates, or an emergency registration of names in the TLD or [deletions] if there are, for example, say we know there is new virus that is using DGA names and we want to turn it off, then we can proceed to add those names under the TLD.

We also get private updates from the EBERO on how the situation is going, and they of course have to do data escrow

deposits. And now let's talk about the specific exercises with it. I should say that in the past we have done, we actually do [practice?] exercises with all of the EBERO providers. We take turns on doing exercises with them, but before this generally, all of the exercises we have done were synthetic.

We were not using a live TLD. However, last year, in September, we received the first request ever from a TLD to terminate their registry agreement with ICANN, per conduct operations that has six month widow and some, that ended March 1st. So we had six months to plan this exercise.

And to make things easier, that TLD had no registrants. This TLD that not actually launch in opening for registrations, and the registry operator was kindly enough to allow to agree, sorry, to agree, to allow ICANN to perform a live TLD exercise with the TLD close to the date when we were going to return the TLD anyway. And nominate one of our EBERO providers agree to perform this exercise, and we scheduled that for 26 January and exactly 12 UTC, to be precise.

Now this is the timing, and Simon please feel free to interrupt me and add anything you feel is relevant to the discussion. This slides are merely timeline [inaudible] when, but let's talk about any other things that we think could be interesting to our audience.

So the emergency threshold for DNS is four hours, according to the gTLD contract. So if we simulated… I should say this is a simulated failure within the TLD did not turn off DNS. We simply said, declared that 12 UTC we are considering that the service is done. So at 16 UTC, all this time you see, at 16:00 we consider that they reached emergency threshold.

So the point we started the normal procedure, and we tried to make this as realistic as possible, trying to avoid giving heads up to most people, we could avoid that so that we could really simulate what would happen in a real situation. So for example, you can see that the next step is the declaration of the EBERO event that's a procedural event within ICANN.

A certain set of people inside have to agree that there is a reality, an issue there that requires the cover of TLD. We take this thing seriously. This is not minutia, we are talking about taking over a TLD. This involves even going to, getting authorization from an ICANN officer. So for example, in this exercise, I can tell you that it so happens that they prefer, let's say, officer that we will go to ask for authorization is [inaudible] the president of the gTLD.

And it so happens that he was travelling to China and had just arrived there. So at the time of the declaration of the event, or when the event will happen, it was passed midnight there. And I call him and he didn't answer. So that's the kind of things that

happen. So I had to go to the next one, the next officer who answered my call and I explained the issue, and he say, okay, go ahead.

And after doing the rest of the process, that's why it took almost two hours to get an EBERO declaration, the EBERO declaration of event enacted. Then we had our colleagues in order to start their process internally, they also have, internally they also have a 7 24 operation. They were very quickly restoring the DNS service, as you can see.

Less than four hours from the, when the emergency threshold was reached, we already had the DNS, DNSSEC, service restore within their servers. So the next thing we did was, of course, send the request to IANA, and here, this was one of the main things we wanted to test in here, because we think the simulated test, the synthetic test that we did before, we had no way, or IANA had no participation in the tests, since there was nothing happening in the root zone.

So we did this for the first time, and to our surprise and pleasure, we found that they were very quickly, all of the actors involved, IANA, NTIA, and VeriSign, were very quickly enacting on the issue, treated it as a real emergency, and the change was done very quickly.

So the total DNS downtime and all of this is simulated, and of course discounting cache effects was 12 hours. So this is counting from when we considered the DNS service down, even though it was not down at 12 UTC, so by midnight UTC, we were having the service already up and running back in the public DNS.

Now we go to the SRS and the RDS timeline. Those were not the services we're failing, but according to the contract if only one of the critical functions is failing, we take over all of the functions in the TLD. Which is, didn't make sense to just take over one, we take the whole thing.

So again, taking the starting time as the 16 UTC when the emergency threshold was taken, and we requested escrow release from the TLD data escrow agent in two hours, and change… And there, for example, we found an issue with them, with that process, that did not allow them the process to be made available within the SLA of the TLD. Sorry. The SLA that data escrow had.

They had some technical difficulties attending our request, and but eventually we received, it wasn't [inaudible] we had there, and so when we had the data escrow deposit, we made that available to [inaudible] and they very quickly restore the EPP service. You can see the times there in the slide.

And then for the RDDS service also, it's restored using the data escrow deposit from the TLD. So the total long time that you can see from the RDDS was similar, two days and three hours. Well actually there, there was [inaudible] time, because the moment we made the change in the root zone, then of course there is no more WHOIS servers from all registry operator working, since all the new TLDs are required to offer WHOIS service within a name under the TLD.

That's WHOIS dot NIC dot the TLD. And finally the last service point of restore, or function is data escrow function, that's with our own data escrow agent. We have a data escrow agent contracted for this purpose, but it also happened in this exercise, data escrow agent that the TLD was using and our data escrow agent were the same entity.

So that's the total exercise, as you can see, lasted for five days and almost six days. And after this, we had the TLD running with [inaudible] until we finalized the, our internal procedures to be able to retire the TLD from the root zone, which that happened in 24 February.

So [inaudible] had the TLD running for almost a month. And I believe that's, oh just one more slide. So this is a summary of the time it took all of the services to be restored. So we think, at the end of this was a successful exercise, we were able to restore

the services. Some was not exactly in the time we want it, it took a little bit more, but it wasn't automatically bad.

And perhaps more importantly in the case of DNS, we'll take a 12 hour turnaround from the time the service first goes down, to the time when we see it back in the public DNS. It's an impressive goal achieved with the help of our EBERO provider. Perhaps the other thing that I haven't mentioned before is the last bullet. During the exercise, we were, say we have a group of people within ICANN that are doing this, and we were taking notes of all the issues we found.

And we also had a [inaudible] providing other issues from their side on what we were seeing, and we find a total of 44 issues. Minor things, procedural things didn't work, people that have lost access to their keys, or access to their servers that had to be reinstated following other procedures. So we have homework to do to improve this processes, but in a year, like I said, we think this was [inaudible] in being able to turn back the service in a LAC TLD in this EBERO simulation.

Simon, would you like to add something?

SIMON MCALLA: Sure, thanks. I mean, I think the, everything you see on the slides is absolutely how it felt from our side as a provider. I think

what was really interesting, I think it was an overwhelming successful test actually, although we had some challenges during the test. I think this was testament to was the fact that the EBERO process was built as a collaborative effort between the EBERO providers and ICANN.

You know, this wasn't… Although we are a contracted party from EBERO, we provide a service as defined by the contract, and you can see all of that in the ICANN website. Actually we built that process together over the course of probably a year, year and a half.

And as we've been doing the tests, versus simulated sort of desktop exercise, and then finally with a live TLD, we've been learning as we go about some of the things that work, some of the things that don't work. And we've been doing that together. I think that's a testimony to that process, and how I think the future of some of these things will be built as the DNS changes and goes on.

What we found, I think, was the stuff that didn't work was really minor stuff. It was silly things like PGP keys not being checked correctly, or escrow not taking longer, as we said, than we thought with the provider. But the fundamental tech and the fundamental basis of the process works really well.

And I think that's great news as we go into potentially around to TLDs, whether they could be thousands more gTLDs, knowing that actually we can fail them out, move them on, and then fail them back again in a relatively short order and successfully with minimal down time. I think is a good thing. So I think, you know, ultimately the program works.

So yeah, positive experience from our perspective. Totally agree. We've got a whole ton of issues that we're now little tiny ones that we're resolving and baking into the next process. I think when we do another test from the other providers, I think it will be even slicker than this one.

EBERHARD LISEE: Thank you very much, and a little bit of an applause is always warranted and appreciated. [APPLAUSE]

PAUL WATERS: Paul Waters. What was the TTL in the DNS record? I'm assuming you did have to change the DNS record for that TLD or not?

FRANCISCO ARIAS: So the TTL, I believe all of the TTLs in the root zone are 48 hours.

UNKNOWN SPEAKER: So that means that the outage that you have could be up to 48 hours?

FRANCISCO ARIAS: Right. That's what I meant when I said not accounting for, I guess I should have said, no, it's there. Cache effects, yeah.

UNKNOWN SPEAKER: Yeah, but now you put it large on your screen to say it's only like three hours and then we were back up and running, but now you're really saying it's like a day.

FRANCISCO ARIAS: Yes, it's true.

UNKNOWN SPEAKER: That's quite a difference.

FRANCISCO ARIAS: It does depend on whether people are pre-fetching, and other effects. And we know that that happens with DS records as well as others.

SIMON MCCALLA: Thanks, just to point out, I think what's important to know here about EBERO is that it's not the same thing as fallback. It's designed for systemic failure. It's defined for failure of the actual company organization, etc. not just saying, we'll provide a backup that will take over if this thing actually technically fails. It more than technical fail, it's organizational fail as well, isn't it?

UNKNOWN SPEAKER: Not strictly. I think the definition varies, but in principle, and we've seen in the past, we've been made ready where there has been potential technical failures or difficulties, and not the failure of organizations. So it could be brought into use if there was a significant systemic technical failure…

FRANCISCO ARIAS: …have access, we ICANN have access to those in a daily basis. So that's the one we use to pass it to nominate, and they quickly restore the answers.

EBERHARD LISEE: Any other questions? Now one question that I have is, are you writing this up? Because some of this might be helpful for ccTLDs to develop their emergency plans and things like this. Not necessarily that we say we want ICANN to do EBERO, but I wouldn't mind, would probably want to have a look at it and set

something up in case we fail, that we push a button and some processes start to make it available again.

SIMON MCCALLA: I mean, this process is specifically designed around transition, as opposed to designed around resilience, and what we've done is, as partners we've worked very closely to try to make that transition process as seamless as possible. So are there learning from it? Potentially, but I think actually, might be better looking at a sort of best practice around resilience and managing that, and might be a better way of accomplishing that goal.

EBERHARD LISEE: Any more questions?

UNKNOWN SPEAKER: So if I understand you correctly, ICANN gets a copy of the zone file for a new gTLD daily. So if you pull the plug on this and run this exercise for a TLD that has domains in it, you may lose the NS information for up to a day's worth of new entries. But would you then pick that up on the backend when you've finally got the actual contact data from the escrow, and fed that all in and did some sort of synchronization or something?

SIMON MCCALLA: So there is a process which does exactly that, which takes a look at the zone and tries to understand whether it's missing records, etc. And then obviously there are various manual processes to try and then recreate those missing records, if that's possible.

But I think we recognize that you cannot create a burden on every registry that has upload the zone file with every change or every hour. It would just becomes technically too complex to manage. So we settled on a day's worth of data and said, you know, accepting that there may be a process that is sort of a clean-up process afterwards for the missing records.

So the kind of thinking behind that on the whole was, by and large, we felt that it was more like to be the more or less active TLDs that were more likely to fail, and therefore that was an acceptable time window. But it is a bunch of compromises, to be fair.

EBERHARD LISEE: In my profession, we call this informed consent. You can get away with a lot as long as the patient knows about the consequences. If it fails, they have one day, you may lose, if you fall in this one day, it's something you need to live with, if you know it beforehand.

THOMAS: You mentioned that EPP and SRS was… Thomas [inaudible] from dot [inaudible]. That EPP was restored, and I'm interested, does the escrow have information about all the registrars about maybe restrictions, AP address, the use of firewalling, because you know, it's usually registrars, they're using registry, they have some IP address blocks that should be set in a registered, to be able to connect for them and all of this information is in escrow, and you restore it potentially, or it's…?

FRANCISCO ARIAS: This is Francisco. No, that information is not in escrow. The escrow is only about the restriction data. In regards to the registers, what we, what the program [inaudible], and is where is a, it's not a hard requirement on the registers, the gTLD registers, but it's a strong suggestion is that they ensure that they are accredited with the EBERO providers, these three EBERO providers, so that in case there is an emergency with the TLD, they can quickly restore their service with an EBERO provider, so they can access the registrations in that TLD.

UNKNOWN SPEAKER: [Inaudible] NIC Chile. Once you restore the EPP transaction, some of them may be new registrations or renewals. What's supposed to happen with the payments associated to those operations?

FRANCISCO ARIAS: So this is Francisco. We, when a TLD is in EBERO, they don't accept new registrations, and there are no expirations, or registrations, ordinarily. Only in emergency situations, we will allow and we will directly request from the provider to add certain domain names, or the list on the domain names, or put them on hold, whatever the proper action will be.

But in a normal course of business, the only operations that are allowed are those that are not [billable?] and that are only necessary to keep up the DNS running. So for example, dates to the DNS, those are okay to do.

EBERHARD LISEE: Okay, thank you very much. That was very interesting, I must say. I appreciate both of you making this presentation. Next will be, if I'm not mistaken, Attila Ozgit. And we'll talk about the recent DDOS attack on dot TR.

ATTILA OZGIT: Hello everybody. My name is Attila Ozgit. I'm the ccTLD manager of dot TR. We had a large, a relatively large DDOS attack on December 14th, starting on the 14th, and then kept going for three weeks, of course in diminishing volume. So I will be telling you about this experience.

Before DDOS attack we were having infrequent small scale DDOS or DOS attacks a few times a year, short durations. Mostly on our registry services, not on name servers. And we were running with six name servers at five different locations, all open source, Linux, Bind, NSD, so reachable at any time.

Average bandwidth, 1.5 megabytes per second per server. And nearly over 1,000 queries per second per server. In the morning of December 14th, basically a DNS amplification attack have started. And towards the end, I mean, starting from the third week, beginning of third week, they changed their target to finance and government sector, went about another week or 10 more days.

And then this appeared, but it was quite a volumes attack, which I will be giving you some numbers later, that we have happened to know from our upstream operators. Botnets sending spoofed query packets to open DNS resolvers and authoritative DNS servers that are not applying rate limiting.

So all these resolvers and authoritative servers are responding to us, to our six name servers as if our servers had created the queries. And 25% of the victims are found to be within the country. Those, all of them, targeting to six name servers. And in the meantime, of course, with low volume of attack, relatively

low volume of attack was targeted to our registry services, web applications.

So it was something like that, from botnet to those open resolvers and authoritative DNS, and from there amplified with the order of 10 to 150 coming to our servers, and varying volumes over the day. However, oops, let me summarize first the communication infrastructure of the country.

Actually three major ISPs are serving the Turkish Internet. Three well-known operators, each connected to tier one at various locations. Unfortunately, on our side, no topology information, and the bandwidth information. But the possible pipes for the attack was through the three major pipes from those three ISPs. So we're calling the ISP A, B, and C. Doesn't matter the names.

However, we were not running with the third ISP. Four name servers are downstream to ISP-A, one is downstream to ISP-B, and one was located at Europe. During the attack, interestingly enough, it was after three days, it was quite harsh in first three days, continuous 24 hours.

After that, it started as a periodic work of daily first shift, between 9 to 17, and the volume of traffic was nearly 200,000 queries per second per server. And reduced rate and different nature of attack was going on during, after 17 hours until next morning. Only name servers were almost always up, however

due to pipelines congestions, big pipe congestion, reachability and delay problems due to overloaded pipes, this is what we had.

Maximum, 200 gigabytes per second attack have been observed in one of the ISP. This is what they have told us. And no synchronized picture of the attack is history. And even I think, those three ISPs are not sharing the information among themselves, to see the whole picture of what has happened in actual three weeks.

And a rough estimation and a press estimation shows that it is one of the largest attacks known so far. Maybe more than 200 gigabytes per second. What we had in the meantime, of course, along with the sleepless nights, make the surface of the attack wider, which means increasing the number of name servers from six to 11.

Two of the 11 are any cast from the help of bind DNS. So effectively, we went up from six to 60 within three weeks. In the meantime, the typical technical work, figure out drop rules to be used, against a varying attack time because attackers were highly adaptive to our defense, once we figure out some pattern of packets and agreed with our upstream operators to drop those packets, and not any longer.

Maybe within 10 minutes, you observe some other type of attack coming in. So it was very well coordinated, as if you are playing a game in the same room with someone else. Currently, infrequently, five, 10 minutes DDOS are still coming in. Of course, we have applied some administrative measures, list of critical domain names, were 100, now it is more than 1,000 this list.

So we are reduced the number of zone updates per day, in the meantime we are trying to make many inspection of zone updates with the probability of that some applications might have been compromised, and people might be trying to do some unauthorized, critical DNS changes.

So we are going back from this mode to the normal mode of operation these days. It was very classical almost all known attack patterns have been shown, however it is not in a mechanical way. I mean, the attackers were sitting in front of some terminals and observing what defenses are being applied, and then changing their attack types.

Both UDP and TCP type of attacks we have observed. I think they have consumed significant amount of resources, paying significant amount of money to the black market. Critical, two critical observations was, knowing egress filtering in the subnets in the country, almost no I should say, including operators. This,

I really wonder if this is a common and known problem internationally, which I mean, which looks to me, is a weird problem.

8%, only 8% of our registered name servers in our database are open resolvers. So that they can easily be used for this purpose, for this type of attack, which is… 8% is something like maybe nearly to 5,000. Another set of observations and lessons, importance of quick root zone management mechanisms is important. We observed that updates were not quick enough, especially on department of commerce checks.

We had to change our infrastructure, maybe I don't exactly remember what, four to six times, not any more than six times, but typically four or five, as I remember. It took us maybe two weeks to do this four changes, especially over the weekends.

This should be, I think, improved. This mechanism should be improved in case of emergency. By the way, this was, I mean, increasing number of DNS servers was with the help of ICANN and IANA, and through their relationship with [inaudible] DNS, and it was started as a quick process but somehow when you come into the [inaudible] web interface, the responses are not coming back as expected.

I mean, it was, or we felt that it was not quick enough, because we were in a rush. What else? Yeah. Effective communication

mechanisms are very important, that could be some important comments for those that might have this problem in the future. Within the registry technical team, be prepared to use near real time communication mechanisms like chatting, jabber type of technology.

Do not rely on email. Any DNS related infrastructure might have been effected during that short period of time, delayed emails and things like this, which forces you to verbal communication, your technical team communicating verbally. And verbal communication could be disastrous for short [inaudible], small times, which we had with the upstream operator the next thing, like your technical team and the upstream operator's technical team are talking over the phone.

It get us for about six hours, for our technical team to prove their capabilities to operators technical team, you know, get rid of the barriers of technical challenges of those people. And then after six hours, the rest was okay, no problem. But initial communication with the technical team of the upstream operator is a problem.

And during the problem, for the rules to be coded, you have to agree on them. And if you ever do this verbal communication through the phone, this is another problem. So you should

always have certain written channels, emergency written channels with your upstream operators.

A funny story, one funny story is, at some point in time, there were some attack patterns which none of our team or the operator team could figure out what it is, how we could manage it. And then finally, operator decided that it is fixed 84 bytes packets, and then let's get rid of them and drop all eight four byte packets.

And then it resulted after six hours in a funny way that data operators, domain name queries, exactly fits into 64 byte packets. And they will become unreachable for a time, for some delta team. So these are the, I think, lessons that in case of emergency, you have to have some agreed ways of communication with your upstream operators.

What else? Effective and concurrent communication is very important during that time. While your technical team is doing a lot of work, you have to talk to IANA and ICANN for root zone management type of things. You have to deal with the organizations, cyber security type of organizations in your country. A lot of administrative work. Handling press and media is another issue.

And effective and concurrent communication examples I've given for upstream operators. And this is it. That's a quick

summary and I would be happy to answer any questions. Not much detail, of course.

EBERHARD LISEE: Thank you very much. That's a very frightening but very, very impressive presentation. And we can all give some applause here. [APPLAUSE]

One question, the bottle neck at IANA was the [inaudible] interface? They were aware?

ATTILA OZGIT: No, the response time… Once you make a query, it goes stages. Maybe eight, 10 stages. I don't know the details of these, each and every single step was what they were doing at the back office. But one of the steps of this approval mechanism is department of commerce check, which takes significantly longer than the other steps.

EBERHARD LISEE: Where you were having an emergency, when you use this other zone management…

ATTILA OZGIT: They know that…

EBERHARD LISEE: My point is, how can, if this happens we speed this up by sort of flagging this, if a request… Just hang on, one thing at a time. That we can reflect this, they know, okay, you're being flooded, you will make changes that if a request comes in, it gets comes through. Jay will have an answer to this because he's on the service level expectation IANA transition. So he can say something, you are next.

STEVEN: Hi. Steven [inaudible] dot AS. First of all, I want to give you my sincere thanks for making this presentation. It's really interesting stuff. I have a couple of questions, I think you pretty much answered the one.

This IANA change stuff has me somewhat concerned. I've had a similar experience, but what I did was exercise the emergency IANA contact number, which proved to be a little rough initially. But once I got hold of the right person, I experienced great service, both within the IANA staff and also with NTIA, and I'm curious to know if you exercised that number and got on the phone with ICANN personnel, and it being a weekend, I'm curious to know what your call center experience was, if in fact, you did that.

ATTILA OZGIT: Okay. We never tried the phone conversation with them. It was all mechanical about it. At the beginning, I have spoken to some people from ICANN management. They diverted me to Kim Davis, and I know him from earlier ICANN meetings, and we wrote, exchange a few emails. And then everything was mechanical with the IANA stuff.

And of course during that time, we have two weekends. And that department of commerce check, unreasonable [inaudible], were not coincidentally of course, relatively related over the weekends. But due to the time difference, from Los Angeles we are 10 hours away, so it makes this inconvenience that could be one reason.

But typical weak changes, took us one and a half day, maybe 36 hours to finalize.

JAY DAILY: This is Jay [inaudible] from dot NZ. Yup. So, I can't tell you what's happening now, but I can tell you a bit more about the plans for post-transition IANA.

So there is, the community is developing the service level agreement for IANA to operate with. And there is a provision in that for an emergency change, to be understood. And of course,

the good part is that the department of commerce will not be involved in future, and so we expect things will be considerably faster because ICANN already well developed processes and know how to do these things, and there is no political decision or other checks that need to be made there.

And then there will be a customer standing committee, which is made up from people like us, who are monitoring the performance of IANA, and if problems appear, then they have power to do something about it. So, I think it would just change naturally because of leaving the department of commerce out, but even if it doesn't get approved, there will be a mechanisms there to do that with.

I have a question, if I can ask though? You said that when you started to drop the packets, the attackers knew that, and then changed their attack. How did they know that?

ATTILA OZGIT: How?

JAY DAILY: Yes. How? How did they know you were dropping packets?

Right.

So but they must have had some means of monitoring systems….

ATTILA OZGIT: Yeah, we are monitoring our own system as well from 10 different probes.

JAY DAILY: Yeah, so there must be some trail back to them through the monitoring. There must be a way of finding them through that. Okay, thanks.

DAN YORK: So it's Dan York. I had a question or a comment to do, but I'll answer Jay on that regard. Jay, you could monitor those servers externally. Somebody else could just see what their availability was on the name servers and then figure out if they come back from the attack and then start up the attack again.

So they can be anywhere out there in the Net, you know, trying to watch what's going on there. So to go to my commentary. Dan York from the Internet Society. I would just say thank you for bringing this presentation. We need more of these from folks who have been through this experience. So thank you for taking the time to do it.

To your question that you ask about are people seeing this lack of ingress filtering and ingress filtering, is that something happening larger? Absolutely. And one of the initiatives that a number of people from within the routing security community, so a different kind of community around here, have put together is something called, it's called MANRS, M-A-N-R-S.

It's the mutually agreed norms for routing security, and you can find it at MANRS dot org. It's manners without the E. But it's a group from within the routing security industry that's looking at, how do network operators come together and agree on what you've talked about? How do you do ingress and egress filtering? How do you block, you know, prevent the IP spoofing that you saw in here?

So I encourage people here from within this community to go take a look at that. And if people are interested, I'll be presenting about that at the technology expert's group meeting on Wednesday, which is open to anybody to participate at. But yes, it's a much broader problem then what you're seeing here.

If we could get network operators from around the world to take some of these steps within their own networks, to make these changes and do what you're asking, to prevent the ingress and egress filtering, if each operator could do that, we would then up

the level of security and reduce these kinds of attacks that are hitting people like you.

EBERHARD LISEE: Okay, we're now scrapped for time, so I don't want to stifle the discussion. Andrew first and then Mark.

ANDREW SULLIVAN: It's Andrew Sullivan. And I work for [dime?]. So I had a peripheral role in this, because people who do work at [dime] now, I just kind of go around the world. But I was involved in some of these discussions, and we were watching this. And to respond to the question about how people are monitoring it, this is not the first time where we have seen back off from drop packets.

And there is a tantalizing evidence which I am not liberty to disclose, that some of the attackers have figured out that they can make themselves a victim as well. So they stand up victim systems, and part of their stream actually is a target. And when they start getting dropped answers at things that they have sent through, what we see is that they notice, oh wait, we're not getting all of the bad answers any more, and that's a sign that, in fact, there may be a problem.

So I think that the attackers have figured out that they can build their own monitoring into their attack traffic, which is not great news, although it's hardly surprising. And I think that that is part of it. And obviously, because DNS is a public service, you can just query the thing and see if you can keep getting answers.

But I think they have automated that.

MARK [ELKENS]: Mark [Elkens] here. Yeah, I would second that. Please give us the full information and mitigation, etc. I'm sure you've probably had very dark hair a few weeks ago. And I know Kenya also suffered from the fact that they couldn't get changes done in IANA perhaps as quickly as they should have been able to.

And Jay, if I can ask a question, is the new… After the contract and the IANA changes, etc. Are we going to be making use of the fact that ICANN is in multiple time zones? I mean, this guy, Turkey, Istanbul where ICANN is also kind of sitting. Are we going to make use?

EBERHARD LISEE: This was better question for Kim Davis, who just came, so we can ask him about it.

He came a little bit late, so I shortly that his name had been mentioned.

ATTILA OZGIT: Can I give a short answer? Unfortunately, I wish I could, sharing information with you fully. But the problem is, you might have it, a bad experience. Your ops stream operators are not synchronized among themselves, they don't know the full picture. You can only verbally take information from them on certain time spots.

So you have very tiny partial information. Not much. I wish I could.

EBERHARD LISEE: Could I take a short interruption and introduce our new CEO, has managed to come and watch what we are doing here. I heard before he is interested in this, so I wanted to take the opportunity to just give him a hand. [APPLAUSE]

If you wish to have a former speaking part, we'll set you up in the next meeting for 20 minutes.

GORAN MARBY: I just want to come in and say hello. Really, nothing else. My schedule, for some reason a lot of people want me to come and

say hello. I wanted to come here, but my staff has pointed me here, because they seem to think you're a very important group, so then I agree. I'm really here to learn. I'm sorry I can't spend more time with you.

I thought it was important for me to come in, so you see my face. Not very much to look at, but that's the way it is. But I'm, you know, if you have any questions, it doesn't relate to what I'm going to do after the end of May, I'm very happy.

I really don't want to disturb you, but please shoot off if there is anything you want to ask me.

EBERHARD LISSE: I, in fact, am good for time. So if there is anything that we want to do, take our time, but I also think it's a little bit of a surprise and I don't want to basically sort of put somebody on the spot. If you…

GORAN MARBY: It's okay by me.

EBERHARD LISSE: If anyone wants to say anything please go ahead, otherwise I appreciate you coming.

GORAN MARBY: Thank you. Can I sit in for a couple of minutes?

EBERHARD LISSE: You can sit as long as you want. Jay Daily has a question.

Jay Daily is the manager of dot NZ's technical side.

JAY DAILY: Hi Goran. Thank you for coming. I gather that you have a technical background, to some degree. I don't know how much, or much about it. It would be nice if you could tell us a little bit about that please.

GORAN MARBY: When I did my first interview after I was, when the announcement came out, a journalist said to me, I'm the nerd-iest CEO from ICANN. I took that as a compliment. Most of you would take that as an insult if someone called me nerdy. I'm a PowerPoint engineer, a white board engineer. And I worked…

I'm not an engineer by training, but I worked with the Internet, IP related Internet since the middle of the 90s. At one point in time, was very good in sounding like a [hayes?] modem, if anyone is old enough to remember them. Yeah, some of you are.

And I'm, I have a tendency to ask a lot of questions when it comes to… With that said, I have a lot of things to learn about

how you do things, but and I'm going to ask you questions a lot about it. So I have a basic understanding. I worked for Cisco. I worked for resellers of Cisco. I worked a lot with information security in the IP world.

So, all my engineers have always said that if it works for me… I'm the perfect pitch of being, of knowing so much that I can be dangerous. I don't think that's a compliment, actually. Was that an answer to your question?

I really don't want to disturb you, so I'm going to sit down for a short while and then I'm going to leave, okay? I just wanted to say hello.

EBERHARD LISEE: You're most welcome. Kim?

KIM DAVIS: Thanks Eberhard. I wasn't in the room when this discussion was happening. Someone tipped me off on Java that I should rush in. So I didn't hear what was actually said with respect to the IANA service. What I will say is just a couple of tips for TLD operators that might have an emergency situation.

Firstly, we have a 24/7 emergency escalation number, if you're not already aware. We send all of the TLD contacts to this

telephone number once a year. That number activates our staff. We have a 24/7 rooster. We have staff that have their phones on, on their day to be on the rooster. And we also have a mechanism of activating our reserve management partners, VeriSign and NTIA.

So that if we get an emergency request in, can rapidly reach them all and have a rapid conclusion to the request. The other tip I would give is beyond calling us on the emergency number, is do some change requests via RZMS.

Many of you are aware, most of the root zone management process is automated today. So if you submit it by the web interface, much of the processing will happen even before we look at it. You send it via email, it will sit in a ticketing system for someone to manually go look at it. That would be the first step.

So if you submit it via our automation system, much of the processing will happen without us even needing to take a look at it. So that's just general suggestions for emergency management, emergency response. I'm happy to answer any more specific questions, but apologies that I didn't hear what was discussed.

EBERHARD LISEE: You mentioned that the bottleneck was the NTIA. But I think maybe a combination of [inaudible] NTR emergency number so that you are aware, you will be getting traffic, somebody must pay attention to this, and somebody must warm up the guys, or wake up the guys at the department of commerce so that they can react quickly, because he said it took, especially with the time zone changes and so on, it took 36 hours to get an emergency name server change.

And that's in [inaudible] just not acceptable. And it's good that he mentioned it, and we all understand that it's reasonable, but we must find, if it happens again, we should learn from this. And I think my view of this would be to warm you up, phone you up on the emergency line we're having this problem, look at this [inaudible], when we make a request on the other, when we phone you on the emergency thing again that you know it's in the system that you can go into the office, if necessary, or access this system and push this through as fast as you can.

So we can then reduce these 36 hours that cost the gray hairs on Attila into something manageable. Already this interaction is helpful now is already helpful I think if it happens again. Warren?

WARREN: So Warren [inaudible], Google. Just to note, 200 gig is a large attack, but it's no longer really a big, big attack. So if people are trying to sort of size names over to deployments for large attacks or really, really large attacks, you know, it's more like four or five 600 gig is where things are going.

And if you follow the curve, you know, mixture it's going to be substantially larger than that. So I guess the sort of good advice is have a backup provider that you can switch to if you need to like [dime] or someone like that for an emergency. That's all.

DAN YORK: Dan York. I wanted to thank you too for highlighting the importance of communications with your partners and the other folks that are there, because I think it's an under-looked, you know, as Warren is raising his hand on that.

I think it's an under-looked part of things. I guess I should ask people here, how many of you have those contacts or the non-email contact mechanisms with your upstream partners and others? How many people here do? A few hands. But yeah, right? That's a critical piece to think about, how do you reach them when email doesn't work?

EBERHARD LISEE: Okay. Next one is Christian Hasselman. And thank you very much, Attila again, that was a very, very interesting and important presentation.

Technically we are a little bit behind, but because we had a cancellation again, we will not have any problems so we can just carry on.

CHRISTIAN HESSELMAN: Thank you Eberhard. So this presentation is about TLD ops, which is an incidence response facility for ccTLDs. So it very much aligns with the previous speaker. This is a presentation that I will also be giving tomorrow at the ccNSO member's day, but I slightly adapted it for this audience.

The technical content is not, you know, it doesn't go that deep in technical terms, so this is more of an overview presentation.

So TLD ops is an incidence response community for and by ccTLDs, as I already mentioned. And it basically brings together those people that are responsible for the overall security of their, security and stability of their ccTLD. And the goal of this community is to facilitate, is to increase the security and stability of the services of individual ccTLDs, and also of the Internet at large, of course.

The approach we are taking is basically twofold. One, we wanted to help ccTLDs to be able to reach each other in terms of, through a contact repository. And I'll be talking about that later on. And we also want to increase the level of shared knowledge of security incidents, security and availability incidents.

So that's in terms of security alerts. Our community is open to every ccTLD, so you do not need to be a member of the ccNSO. Everyone can join. And the community is being governed by the TLD ops standing community, which consists of five ccTLD representatives, plus three people, three liaisons, one liaison for SSAC, one for IANA, and one for ICANN's security team.

The TLD ops community revolves around the TLD ops mailing list. And it's basically, the mailing list basically serves two purposes. One is to act as a [contact] repository that enables ccTLDs to quickly contact each other so they can, you know, contact each other for help or for sharing information, perhaps such as in the event of the attack on dot TR.

And it works as follows. The mailing list, once you're subscribed, you get a regular automated email, and it contains the contact information of all of the peers, of all your peers that have been subscribed to the list, so that's contact person, phone number, and also email addresses.

And the idea behind that email is that you can store it in your inbox, so that you can also look up the contact information of your peers in offline situations. So for instance, when your DNS services are no longer available. And the other purpose, in this email, there is all of the contact information that's, of your peers in there.

So everyone who has subscribed to the mailing list is in that one email you get every month. And the second purpose of the mailing list is to share security alerts. For instance, alerts on mail redistributions, software vulnerabilities, DGA algorithms and that sort of thing.

The thing that set it a little bit aside from similar initiatives in this, in the incident response field, I think is that we chose to use personal information on, that members need to subscribe with personal information. So this means personal email address and also their first name and last name and phone number. And that we do not allow role based accounts. And the reasoning behind that is that we think, we believe that this will contribute to further increasing the level of trust within the ccTLD community because people start recognizing each other's names, for example.

Also what's a little bit of different is that we do not use the common vouching model in this community. We use a model in

which ccTLDs are being admitted to the list through approval of their IANA administrative contact. So if you are not subscribed to the mailing list yet, you ask your IANA admin contact to send an email to the ccNSO secretariat, in this case, saying that, okay, these two or three folks are the security and stability people of my ccTLD, and then we'll subscribe you to the list.

Or alternatively, you can also send an email if you're not the IANA admin contact, and then CC the IANA admin contact. So more details are available on our website, and I'll show the link later on.

The rationale behind that decision is that we think that this provides a low entrance barrier so that we can get as many ccTLDs on the list as possible, which is unlike the vouching model which requires you to, if you want to join you need to, other people need to know you before you can actually join a list that works based on that model.

So these are our current membership statistics. Overall we have 173 ccTLDs on the list, so that's 60% of the total number of TLDs. And if you look at the ASCII ccTLDs, we currently have almost every ccTLD from the European region on it, and North America is actually advancing quite well too with 60%, and the regions where we have some work to do is in Africa and the Asia Pacific and in the LAC region.

EBERHARD LISEE: Okay. Oops.

Then we'll take this as a very, where does [inaudible]…

CHRISTIAN HESSELMAN: Not really because…

EBERHARD LISEE: Make it without the slide.

CHRISTIAN HESSELMAN: Okay, I'm not sure what happened there.

Okay, some words on what we did. In the past couple of months, we added six new ccTLDs. We updated the leaflet that we developed a couple of months ago, so you can download it from our site. The URL is there. And we also extended the TLD op standing committee so that's the group of people that governs the whole community with, for instance, a few liaisons to SSAC and ICANN security team, for instance.

And we're also working with the ICANN regional vice presidents to onboard more ccTLDs. Our work right now has basically two challenges. One is how do we enable ccTLDs to share security

incidents on the list, or security alerts. What we want to do there is we're, first of all, we're working with ICANN, so that they can…

We're working with ICANN for, to enable them to share security incidents that they are seeing on our list, with the final objective of actually having all ccTLDs sharing such information on the TLD ops list. So that's like a peer to peer form of communication.

And our second challenge, second part of this challenge is to, you know, get somehow get incorporated into the operational processes of ccTLDs. For instance, to end up on check lists and that sort of thing. We really don't have a good solution for that at this point, so if you have any suggestions then please come and talk to me, and I'll be glad to hear them.

Our second challenge is to perhaps include gTLDs. Of course, this whole concept might also work for other TLDs, not just CCs. This might however, change the nature of this initiative a little bit, because of course, there is many more gTLDs than there are ccTLDs. It would, however, strengthen the incident response capabilities of all TLDs and not just the CCs. So I think that would be very good for the Internet and for the stability and security of it.

So this is something that we would be discussing tomorrow morning with the ccNSO community, and we'll get a feel for the

room temperature there in terms of what they think about gTLDs joining this initiative.

So our planning until the next ICANN meeting is to further increase the number of subscribers, of course, because that's important. We also want to get more information to be shared on the list, security alerts. That's why we're working with ICANN security team, and we also want to develop a simple procedure for members to enable them to update their contact information should that be necessary, because we only have 320 people on the list, so there is bound to be some folks that leave the company, or maybe change roles within their organization.

And finally, we also need to, we're also looking for representative from Asia Pacific region to join the TLD ops standing committee, because we have a vacancy there. So this is our URL. These are the folks in the TLD op standing committee, and if you have any questions, then I would be very glad to take them.

EBERHARD LISEE: Okay, thank you very much. A little bit of applause is always appreciated. [APPLAUSE]

To start the discussion from the Chair, I don't subscribe to this because when I suggested that my [inaudible] generic email

address that we are using for these things, and that is guaranteed to be answered, even more guaranteed than my personal email address, it was not accepted by the secretariat.

So as long as the secretariat tells me which address I must subscribe and which address I cannot subscribe, this is not something that I find of much value.

CHRISTIAN HESSELMAN: It's not the secretariat that tells you that, it's the standing committee that sets this policy.

EBERHARD LISEE: Yes, but this is the policy that you want to have, this is the policy…

CHRISTIAN HESSELMAN: I know your position Eberhard, so yeah.

EBERHARD LISEE: I just want to make it, put it out there, maybe as I said, want to have the generic addresses, I can guarantee that my generic address is being read. I cannot guarantee that my own address is being read all of the time. So maybe you want to look at this, how you deal with this.

CHRISTIAN HESSELMAN: Well first of all, we ask members to subscribe with at least two or three people, so there is some redundancy there. I try to motivate why we did this, and this is something that I would like to stay [inaudible] with in the near future.

EBERHARD LISEE: Okay. Any other questions? Andrew?

ANDREW SULLIVAN: This isn't a question… This is Andrew Sullivan. This isn't a question but a suggestion for that. You could ask that everyone who is being a generic address have also subscribe with maybe no mail so that they could all post, but it all be read by a generic mail, thereby solving the problem that Eberhard claimed, which is that, you know, maybe it's not being read all of the time.

So you could get the benefits in both directions. Just a suggestion.

CHRISTIAN HESSELMAN: Thanks.

EBERHARD LISEE: That would do it for me.

Anybody else? Warren has a question.

WARREN: So Warren [inaudible], and also part of the TLD standing ops committee, so replying to Andrew. There is actually a lot of research that shows, especially from sort of the security trust groups, that roll accounts end up going to folders and ticketing systems that people just never look at. So you know, having it subscribed from a roll account or a search generic account, in many cases, means it would never be read.

And I don't know where the paper is, I was trying to find it. But it's actually, you know, useful research that shows this, not just randomly made up stuff.

EBERHARD LISEE: As I said, my roll account is guaranteed to be read. That's my norm. And then I think we close the queue and Jaap will do the next one.

[STORM RICHIE]: [Storm Richie] from Secure the Name Foundation. Quick question for clarity. So as a new cyber to this list, can I send an email to this list and it hits all of the ccTLDs? Or is this only for,

or all TLDs, or is this all for TLDs communicate amongst themselves?

CHRISTIAN HESSELMAN: Right now, it's about ccTLDs communicating amongst themselves, but we also have John Crain and Kim Davis on the list for sharing security alerts that come from the ICANN side, so to speak. So we have one external party, if you will, that can also share security incidents on the list.

[STORM RICHE]: Okay, but as a security researcher, what would I do then? Would I go to ICANN? Or would I send it to you directly, or someone I know on the list?

CHRISTIAN HESSELMAN: That's a good point. To be honest, we haven't really thought that through. So that's a very good question.

[STORM RICHE]: Okay, so that would be great for the next meeting, I guess.

CHRISTIAN HESSELMAN: More objectives, oh my God.

EBERHARD LISEE: Okay, very last question from Jacques.

JACQUES LATOUR: Jacques Latour with dot CA. So one thing to remember is that, twice a month we send the entire list of contacts and phone numbers and all of the information to all the members of the mailing list. So that's not information that we want to get sucked into electronic system, ticketing system. That's something we want to stay within the TLD ops community.

So that's one I expect we need to think about when introducing new members to this list.

EBERHARD LISEE: Thank you very much. Next will be Jaap. We're still waiting for the presenter from IFC who has indicated he is on the way, and he will wave to me when he enters the room because I don't know him.

Hasn't waved yet.

JAAP AKKERHUIS: Let's hope I can find the right button later.

This is just like it is, okay sorry.

So, what is this about? Well, it's all his fault. I was asking him a question and he said, well, how do you know this? And I told him the details, and now I'm going to explain all of the details to you as well.

EBERHARD LISEE: It's an old management technique. When you are asked a question, assign a task.

JAAP AKKERHUIS: Right. So, what's the bottom, you said it was?

Doesn't really… Oh, that person. There are too many patterns. I really want to go back to one pattern marks. Anyway. Well, why do we want to do monitoring? Well, this monitoring of DNSSEC, of how the top level domains actually maintain their DNSSEC stages, that's what this is about.

Why do we want to do monitoring? Now, DNSSEC needs maintenance. It's not an install and forget application, as it used to be. I mean, it used to be just [inaudible] on some [inaudible] which was [inaudible] not using it, and that was what you used for your authoritative server, and that's it.

And why are we doing this? Well, I'm just curious. I wanted to know how well this is done, and especially since Walter and I

kind of, sometimes get questions, oh your software stinks, you look what it is and it's actually complicated. On the other side, and so it's easy to know about that things in front.

What do we check? Well, just TLDs. We just let's [inaudible] to a small stride, and it's not really going to be a big project. And how we're doing? Well, pretty easy and pretty cheap, because we don't have the time, and we don't have the money to do it on a real scale.

So it was just done, I think it took an hour to do this, to set this up for the first time, and so what do we do? We all, first find all the signed DNSSEC signed TLDs. So with this line, you actually find, kept the root zone from somewhere, a couple of places where you can find.

And some root servers has it's, ICANN actually has two different servers, which is especially meant for [inaudible] the root zone. So there are a lot of places where you can get it, and if you have it, if you just look whether or not DNSSEC [inaudible] and then you create from that, all the list from all the signed root zones. And that's it.

And then the second part of this is how to file a date with these domains. Well, we just happen to have a small little program [inaudible] which you can, which actually do all of this validation for you, if you want to. So you ask for the [inaudible] record of

the particularly a TLD, because this always [inaudible] delegation, or else it won't be delegated, right?

And if [inaudible]… and then you absolute do it as well. And so, you scheduled by this by [u-con?] on your Unix box, I do it twice a day and that's it. And so, it's a [inaudible] of their lines again.

Well, what do we get? We started to do this three years ago, and I got about 500 mails, actually it's one more, I got one today as well. And there is a lot of repeats, and some things are stuck forever. I mean, there is one ccTLD, we've got four emails, and then I decided to, I got enough of it so I actually changed script and have a stop list.

So I have stopped this ccTLD, sometimes I look whether or not it's back online. And yes, it's back online and so are able to stop again. And that's not a big deal. And [inaudible] well I will come back to that later. What I get is about 30 different TLDs in these [inaudible], which actually have problems. And the interesting thing is if you compare this to whether these are new generic TLDs or ccTLDs are all generic TLDs, most of it is actually ccTLDs.

It's kind of surprising, but or it may be not because a lot of the ccTLDs don't have a lot of money so they might have problems there. And this is the whole [inaudible], these are all of the things I found, minus one, I forgot dot [inaudible], that was just one Sunday morning. These are the whole change.

These are all of the TLDs I find problems. Some of them I just found, as I said, more, offered about a year period, and it's actually Mongolia but IDN type of Mongolia, but they didn't have any delegations on it, so who cares? That's probably why it took so long before they find any notice.

And so these are, some of them I just see popping up, problems. And then they actually carry fixed, some of the new TLDs you see listed here, like cancer research, feedback [inaudible] just… Yeah, these were actually for one day, and then all 12 of them, and then they disappeared. And if you look to a lot closer to these things, they share the same interest.

They have the same names of, it looks like they have the same name services after. So there might be something [inaudible] at the moment of getting into production, or their software might break, and so all of them, they share a common fate. It's not a big deal. I mean, some of them are pretty desired [inaudible] dropped out from the world, at least for DNSSEC one day. It's kind of interesting to see.

But after if you see, after a three year period these are… Some are actually repeats. Some are repeats and by every month you see them popping up and then they go down again. Well, most of the time it's expired, 62 [inaudible] somebody is sleeping.

[Inaudible] have apparently [inaudible] problems, me and… IDN for me and [inaudible] properly, popping up for a while and then after a couple of days, it just be here again. And then it goes well for a couple of months, and then suddenly it pops up on the and it goes back again. So but actually it's surprisingly stable what you see.

What the error types, expired 62s. I mean, this is what they see a lot, oops. I mean somebody was sleeping and didn't do his job, or forgot his current job [inaudible] the 62. [Inaudible] but complicated what you see often, is no matching algorithms. I mean, [inaudible] to the new key, but forget to put it after the root or the other way around. And but these are here most, be gone after two days. Something like that.

So that's… And then suddenly people forget to sign the zone, you know, this is kind, which I think, well sloppy operations. They are not consistent way of doing business. And I'm spending a lot of server fails, which can be anything, can be server fail. The guy went out in, network problems, you know, I mean, most of them, these are short living anyway. So it's not…

That's basically the roll around what we have. If you look [inaudible] a little bit by TLD type, most of the ccTLD which have repeat failures, it's the wrong, the top. gTLDs, well it's kind of

fail most time, and it might be [inaudible] problems, it might be soft problems. I mean, this is, as I said, very cheap.

And it will take a lot of time, and I really don't go into looking at every time. I see this popping up. I mean, and only when it take a long time like part one, I had some problems for a long time. And that's when I asked my neighbor whether or not he could give somebody a kick of that. And then new TLDs which I [inaudible] actually, single failures and start up with shared, because if you see 12 popping up today at the same day, and all 12 are actually bogus, and then and this be [inaudible]…

That's my guess. What now? Well, [inaudible] is pretty easy. It's just [inaudible] and actually some other [inaudible] this way [inaudible] from [ripe?] used to be another [shelf keeper?] send me email until I showed it to Daniel. Said, oh, and then well that was the start of DNS mode, and it's now a lot of people lining up.

So it's also that longtime monitoring might actually show a possible trend in behavior of the operations. I mean, what you, a lot of times these people are used to in short, I mean, just one offer like the data stuff once a year, which only gives you a snapshot and you get all of the sample errors. But during this consistently for a couple of years, you might publish [inaudible] and now, we're actually whether I should turn this into surface

for other people [inaudible] so I can watch their domains and make this slightly more robust and then just now.

Things like that. Well, if there is interest, just let me know. And… There are actually other DNSSEC tools. I mean, this what is there [inaudible] early warning system which used the website. And if you looked to that, it will tell you the things like, well, you [inaudible] to expire within the next seven days, or within the next three days, and it's pretty slow, and you have to go to the website to find this. Even though that's an awful lot, I have seven days, about 20 or 30 you see popping up there.

Then there are all the other DNSSEC checkers, which are useful especially after the fact that things are down to figure out what is really going on in details, the DNSSEC [inaudible]… more official DNS face from doing this, I forgot who was doing that. So I can [inaudible]… is also doing this stuff for you.

And this is actually based… Any questions?

EBERHARD LISEE: Thank you very much. A little bit of applause first. [APPLAUSE]

I always like these small little quick and dirty jobs that you can run in a share script, because I can manage to do it in my own system then, for example. The last show DNS with a zone master and a DNS checker, and the last zone master and DNS, they have

presented zone master, I think, in Dublin and DNS in LA, if I'm not mistaken.

Yeah, these are web based things for which you can also get per front or backend or whatever. Any questions?

JAAP AKKERHUIS: There is always, all of…

UNKNOWN SPEAKER: Yeah, can you change the output of your very useful script to be Twitter feed?

JAAP AKKERHUIS: To be what?

UNKNOWN SPEAKER: A Twitter feed?

JAAP AKKERHUIS: I don't get your last words.

EBERHARD LISEE: Twitter…

JAAP AKKERHUIS: Oh, Twitter. I could do that, yeah. Why not? If people want that. I have to make it slightly more [inaudible] because sometimes because of network failures, I can't care to use zone file. And we didn't want to see that popping up on Twitter.

EBERHARD LISEE: Actually I don't think you want to see any of this pop up on Twitter, but not because you don't want to see about Twitter, because you don't want it to happen. I also think the reason why ccTLDs get caught more often is because the gTLDs are in a much more stress and control, and if it doesn't work, they get a [inaudible] and finishing the [inaudible].

So if you have three or four start up at the same time, it's obviously that they're on the same infrastructure, but if it's only for one day and it gets fixed, it's not a problem. I must say, I like this little tool so you can look at these things. I have been, as you know, trying very hard to get in touch with the guys at dot BW. I even met a colleague here at the meeting. I have communicated with the colleague, I'm not even mentioning the gender.

And I was promised that I would be communicated back which hasn't happened yet. This is the way sometimes things happen with us in Africa, I must confess.

DAN YORK: Dan York. Just thank you Jaap for making this slides and the information available. I guess a question for you, will these slides be posted up to the tech day site sometime in the next week? Or are they there now?

EBERHARD LISEE: All slides will be, are or will be posted to posted to the tech site. On every slide, we don't take presentations and do not post them.

Okay, Jaap we will give you the email address from the up TLD site mailing list for dot BW so you can communicate with them.

JAAP AKKERHUIS: Okay. So your completely different. Now I'm here and we're still waiting for the [inaudible], I have three [inaudible] with me, talking about [probes?]. So if there is some interest in [inaudible], come and find me.

EBERHARD LISEE: Has Brian arrived? He mentioned that he was having taxi trouble, so we will have him very ably substituted and represented by Vicky Reed, Vicky Risk.

Should he come in while you're doing, I will have him substitute you in mid-sentence.

VICKY RISK: Good morning. I'm the backup presenter for Jeff Osborn. He is unfortunately in his hotel room this morning wishing he were dead. So these slides are actually much more beautiful than those that I usually make. So I'm going to talk a little bit about what we've been thinking about IFC, about the future of F-root.

As most of you know, we operate one of the 13…

EBERHARD LISEE: Can you turn down a little bit? Hello? Jeff, it's interfering actually.

VICKY RISK: And oversee one of the 13 DNS root nodes. So this is the agenda. We sort of, took sort of a philosophical look at what is the purpose of the root? How did we get to where we are today with the root operations? In an effort to think that there is a new way that we can approach operating the F root node, that is perhaps a lighter touch, lower maintenance model, that is more efficient to serve the current purpose of the root.

Presumably this audience is pretty well aware of what a root server is, what they do. A root server is really just an IP address, and someone with that address agrees to maintain the data and to answer the queries. And of course, these days all of the root nodes are any cast.

Traditionally, there were exactly 13 devices, 13 physical servers that would answer root queries. And so each one was 7.7% of the world root server capacity. Most of those servers were in the US, and the failure of any one node of course, especially a node outside the US, could have a significant impact on root operations.

So these root servers were built for very high availability. Today, there are at least 572 root server devices. This is as of a month ago. So if one fails, or is taken off line by attack, it is a very small impact to the global capacity to serve the root. So it is not necessary for every root server node to be built to space shuttle specifications.

Individual root servers have instead acquired a new role. We just heard a presentation about a massive DDOS attack against dot TR. It's very useful in case of an attack of a root system to have some sacrificial nodes. Of course, the root server still has to serve the root, but it appears that it may be the case that instead

of having very high availability nodes, there is also room for having sacrificial nodes that will absorb the attack.

Historically, when we look at root server attacks, we look at the failed servers. But given the current numbers, perhaps it's better to look at which servers did not fail. How many were left standing? Even in the attack last October, many more than 13 were left standing.

So again, individual servers don't have to be bomb proof.

So our current thinking is for ISEs, root server in particular, that perhaps we might focus on having our nodes closer to the edge of the network, where they can intercept attack traffic sooner, and closer to the origin of that traffic. So obviously one of the purposes of root servers is to give fast response times.

Up until today, our strategy for deployment of our F root nodes has been to focus on putting them into inter-exchange points. Inter-exchange points that were near a population center, a large number of Internet users. So what we're thinking here is a slightly different strategy of putting them closer to the edge of the network.

One of the issues with maintaining the F-root service nodes around the world is maintenance. I don't have the picture of this, but recently one of our operations guys sent around a little

note. He asked one of the node operators to give him a picture of what was on the consol part, because he was troubleshoot something. And what he got was a photograph of the consol port of the back of the machine, with the cable coming out of it, and this is…

There are communication problems like this. My understanding, and I am not on the operations side of this, but that this is a significant challenge to maintaining this high availability system that you're communicating with a number of people. Some of them may not even really be aware of what the F-root or a DNS root server does. And you're trying to maintain it on a continuous basis.

So the thinking is, well, if we had 1,000 of these, if some of them went down, we wouldn't have to manage them so actively. It would not be a problem. If we had a large number, we could simply take a role, and see how many were up.

So Brian is here, he can take the questions. So right now, we do have a number of F-root nodes out there. We do have 50 instances in 50 countries. Most of them fill a rack. Most of them have a dual route servers, dual servers, separate consol port machine. So they're a big nodes that are designed, of course, for high availability.

And they are actively managed. Want to take over? Okay. So, what we're thinking about is what smaller root servers look like. And I just want to mention that these ideas, these concepts, these are not just things we dream about in isolation. A lot of people have helped and contributed ideas.

So one idea that, we've talked about this for quite a while, is a Dell based F-single. Just a single when you rack mounted server, replacing what now is five pieces of hardware. Another idea is a really small…

Oh okay. So one of those F singles just went into service about three hours ago. Another concept that we're working on is a small form factor standalone server. There is the Beagle, the Minnow, and I'd never even heard of the Pine, but anyway, they all have their advocates within IFC. Different people have their preferred little tiny, tiny root server box.

Another idea is docker container. Obviously this is something that doesn't require hardware installs, so that's a value in places where, that's extra overhead. And the last idea is the RFC 7706 which is running a root server on loop back, which is on [inaudible] draft.

So what we're thinking is to make all four of these options available.

This is for questions, and this is Brian Reed who just walked in the door, and these are actually his slides.

EBERHARD LISEE: Thank you very much. And even a bigger applause for doing this off the cuff. [APPLAUSE]

And he blindly not only walked into the door, he drove in with a taxi and got off at the wrong place, so he was actually really lucky to make it in time. Any questions? I have a question. When am I getting ours?

BRIAN REED: Give us a mailing address.

VICKY RISK: Which one would you like? Are you looking for software only…?

EBERHARD LISEE: No, no, no. I'm just joking. Basically, we can, I was looking at the F one factors, that's more one new thing that we can talk about. I mean, I'm just saying, the point is not when I get mine in dot NA, I'm getting, I'm saying it becomes now so easy and so simple, we should basically start queuing and tell them I'm ordering them.

VICKY RISK: So I think that F single was actually specifically announced at AfriNIC last year, because part of the intended target were in fact the African, Africa locations that needed a lower cost.

BRIAN REED: But it works well enough for anybody. We are startled by how well the F single works. The deployment, we drop ship a box that contains a Dell computer configured according to our standards. We send them instructions for them to make three Ethernet connections, and then phone us and let us know it's ready, and we configure it remotely and turn it on, and everybody started choosing and that's the end. It's nice.

EBERHARD LISEE: Paul Waters next.

PAUL WATERS: So are you going to add a default F-Root configuration in Bind?

BRIAN REED: Bind is for all roots, not just F-root. I mean, there certainly is. We've had it for years a configuration for Bind that makes it behave like a F-root. But although we want to proliferate F-root, we don't want chaos. We don't want people bringing up their

own F-root without letting us know. And if we made the code widely available, then who knows where there would be rogue F-roots?

I personally don't think that's a problem. Here at an ICANN meeting you have to act as though you think that's a problem.

EBERHARD LISEE: Any other questions? All right. Thank you very much. We are exactly on time. I'm quite amazed with how exactly on time we are. I'll see exactly in an hour. I see you're exactly at 22:02 for the next two presentations and the roundtable. Thank you very much. [APPLAUSE]

Come sit down, sit down, sit down please. Okay, good afternoon. As usual, the ones that are here get punished for the ones that come late. That's the usual way of doing this.

We'll start the proceedings with Don Hallender and his team, about universal acceptance. Our colleagues or however you want. Team is good. There you go, go ahead.

DON HALLENDER: [Inaudible] Eberhard. So I wanted to talk, there will be three of us, there is certainly two, and we will multiply materially shortly, about a geek's guide to universal acceptance.

So this is today's agenda, what is universal acceptance and why is it a geek's issue? We'll do a couple of case studies, talk about how geeks can help. We'll introduce you to the UASG. Review quickly some cases and quick guides that got published this week. And we'll talk about EAI. We heard from Marvin earlier today about EAI. And why is it so hard?

And then we'll talk about accelerating some solutions to how, what we're doing and what we would like you to help with in terms of making things happen faster. So, universal acceptance issues that have been around since 2001, when the first of the new gTLDs got introduced. And so that's the dot info, dot biz, dot Asia, all sorts of time space.

It became more of an issue when the IDN ccTLDs came in, and it became an even bigger issue when the, all the new G's started coming in because while the DNS itself worked and the names resolved, the applications that use domain names and email addresses didn't take kindly to these new TLDs.

So we spent, so the issue has been around for a while, with some anger since 2006, 2010, so six years and 12 months ago, the group from the community got together and said, you know, we really should do something about this, and we formed the UASG. And we spent the past year or so getting ourselves in order, getting our ducks aligned, and getting ready to move forward.

So one of the things we did is we came up with a definition of universal acceptance, and it's all about accepting, validating, storing, processing and displaying correctly and consistently all Internet enabled applications, devices, and systems. So having a definition is a key point.

So why is this a geek's issue? The solution is distributed. Almost every Internet focused application needs to be adjusted. Think about Y2K with an Internet focus, except we don't have a deadline. The problem is software more than hardware, and the problem is not generally a routing issue, it is application software, counting system software, web application software sort of things.

It's not difficult, except for the EAI stuff, which is challenging, but it is effortful. So if you're a CIO in an organization, the big cost of fixing anything is opening up the hood or the bonnet, that's the cost and the big cost of opening it up and closing it, and doing the work on the engine is not so bad.

So while it's not difficult, it is effortful but there are some hard bits. So the hard bits are around EAI, email address internationalization. So this is, there is two bits of it. So IDNS on the, I mean, the domain name, or not so hard. It doesn't… It would be better if they displayed properly, but it's not usually a problem.

It's where you have Unicode in the local part, so the user name, the stuff that's to the left of the at sign in right to left scripts, left to right scripts, sorry. The right to left scripts, so Arabic and Hebrew, is really hard. And the normal challenges that you get with Arabic scripts and other scripts that have letters that change shape, letters that are homographs for in similar languages, those are real challenges.

And if you have an email address that's Arabic at Arabic dot Arabic dot Arabic at Arabic, which is the local part and which is the domain name? Unicode issues are a challenge, and the version of IDNA that's used in and included in your software are hard.

There are still a couple of unresolved characters in that space, and we have an evolving world. And I feel like I'm just reading the slides, so let me just talk on behalf of ICANN IT. And anybody from ICANN IT here? So I can tell the stories whether they're true or not.

So ICANN obviously has a keen interest in getting their systems UA. We've been getting reports from Ashwin who is there CIO at each of the meetings since last year, which has been very useful because it has provided an ongoing case study as to how a CIO or how a IT department will go about addressing the UA issues.

So the first thing they did was they had to create an inventory of all of their systems. And while some of you might have really excellent documentation processes within your IT shops, most don't, and ICANN was not perfect. So the first thing they had to do was find all of the applications that they use.

And then, they needed to document what they had. And then they divided it into a couple of different bits. So things that ICANN should be able to fix because it has the source code. But can it find the source code? Do they still have people who can work the source code? What it can't fix, so things that use software applications that they buy in, their accounting system, their email system, things where ICANN is really a little fish in a big pond.

Things that they can get others to fix. This is where ICANN is sort of big fish in a little pond. But then that raises the whole issues around contract. So if you go to your vendor and say, "We'd like you to make this UA ready," the vendor, account manager will start smelling money. And then…

RICHARD MERDINGER: Pardon me, I just want to interject. I'm Rich with Go Daddy. UA ready is not an additional feature, it should be an intrinsic capability in the software. So it just irritates me when people treat it like a feature.

DON HALLANDER: So if I'm a software supplier and my software works perfectly well now, and you want me to add extra features, then that's a feature.

RICHARD MERDINGER: I understand.

DON HALLANDER: So ICANN IT found 84 business solutions, about half were off the shelf. For 19% of them, UA was not an issue because they didn't use domain names or email addresses anywhere in the system. And 80% needed adjustment by supplier. Half were customized or in house, that is they were fixable. And they've gone through and looked at the systems and said, "Yup." So that are relatively easy, some don't need to be fixed at all.

And they reckoned that it would take 18 to 24 months to make changes. So when we heard from Ashwin yesterday in his regular update, he went out to his user community and said, "We'd like to fix this application because it doesn't work with these domain names, new TLDs." They said, "But it works perfectly fine for us. Don't you touch it, because we don't need to go through that aggravation."

So this is real world, real life situation where a CIO goes to his client, and his client sees no value in making the change. So he has now revised his estimate from being able to fix the things that are fixable from 18 to 24 months to somewhere between 24 months and a long time after that.

So what I'm going to do now is Rich Merdinger and Dennis Tanaka from… Rich is from Go Daddy and Dennis is from VeriSign, have been working on an Internet industry review. And so he's going to do that, and I'll press the button for you.

RICHARD MERDINGER: Thank you Don. And the reason that we started doing an analysis at Go Daddy and VeriSign is because we represent pretty much the standard situation that exists in the industry where you have a provider and you have a client that has systems that interact. The wholesaler which is VeriSign is, gets to their ultimate customers through us, so they have needs and requirements at the registry level.

We have needs at the registrar level, and the systems that we have are interrelated, but they're also very different. And that's not all that different from the airline industry, or any other types of industry out there where there are suppliers trying to deal with retailers and customers, etc.

So we know this industry so we started with it. It's not that they're all that special. Go ahead. So essentially it comes down to eating your own dog food. We are not only a set of companies that work within and support the domain name and Internet infrastructure. We actually provide the products that we're seeing that may or may not be working properly in the ecosystem.

So we better make sure that our systems are robust enough to handle them. So eating our own dog food. In order to do that, and many of you may be familiar with this, the idea of getting into a room this size, and if we were to go around the room and say everyone, now, tell us about all the ways that you're systems are broken, and are inadequate, and aren't able to handle something, no one is going to raise their hand because they feel like they've got to project the image of their company so that they can take care of their customers, etc.

So we took this approach as being under NDA, VeriSign and Go Daddy got together with a small group of people, including our internal technical experts for the systems we were looking at, and considered this really more of a joint project that we were doing. The idea being that we weren't going to come out with a blueprint for the registry registrar world on everything that's wrong and everything that needs to be fixed.

Rather we would identify those things within our systems, and then extrapolate those out to topics that could be applied to any industry. The types of issues that we found are probably similar to issues that you will find in other areas.

And to do that, you couldn't now, but to that now, the approach we took was to look at the data model that we use… You can go ahead to the next slide, I guess. We took a look at the data model that is in place that defines the registry and the TLD and domain construct that is used in the EPP system between the registry and the registrar.

We found that the EPP system is always very, already very robust when it comes to dealing with different character sets, and email addresses, etc. It basically serves as a container for where the problem lies. And the problem is really more in the way that we are… The business use and the business logic that is tied around the domain content, whether it's the domain name length, the domain name character sets, etc.

So the EPP system was very, very good. It came down to a few uses of terms such as host names and labels within there. We as an industry have been doing it for a while. When we talk about passing data back and forth, we'll say pass me the domain name. Well it used to be enough to expect that it would be a series of labels separated by a dot with a few possible variations.

Now the question is, are you passing me an A label or a U label? What should I expect? And what we've decided is, at least for our best practice, is it's most effective for the system to be required to expressly state which form that the domain name or the host name is going to be specified in, and if you have a scenario where the name is passed without an explicit designation, that it would fall back to some specified default.

So you end up with behavior that's predictable and defined going in. Another brief topic here is just, we talk a lot about domain names themselves being a problem. But the proliferation of domain names hitting the root recently means that we are now dealing larger numbers of domain names, and registry systems have become segregated into, some of them into farms of SRSs. So as a registrar, you may need to connect to, instead of one SRS to allocate all of the domain names, you may have configurations in place now where you have to proxy out to figure out which one to go to.

So there were some complexities that came along with just even the system configuration as a whole that we didn't expect. So again, as I said, don't make assumptions about… Data integrity is one of the most important pieces. Don't make assumptions about the data, be explicit to what you are passing and what it means.

An example could be, in Don's which side of the at sign is the local part? You should have defined set of rules in place between you and your vendors so you don't have to make assumptions and potentially get it wrong. Also when you do have your implementation, assume the other person on the other end isn't going to follow the rules, so have robust implementations that handle exceptions, raise them in a controlled and expected manner, and can lead to more predictable results so that your customers don't have a negative impact, and just find themselves waiting for tech support to solve something.

The customers themselves, as much as we love to service them, are really a large part of the problem in this case, because they're the ones that are going to buy into the new gTLDs that we're selling, or get IDN names, or use EAI email addresses. That if we want to service our customers, we need to be able to support them. When you look at the domain name industry, we're in an interesting situation right now in that we are allowing customers to provide us, we're requiring customers to provide us contact information.

Our systems and specification only supports essentially ASCII at this point, yet we're in a transitionary phase where we're supporting more international domain names. We're looking at transliteration of WHOIS data as a possibility, and we're semi-

embracing UTF-8 as a means of expressing content, but the underlying protocols haven't been upgraded yet formally to support that, at least not in the gTLD space.

So we are finding ourselves needing to interpret or homogenize data sometimes that comes in, so it's in a way that our systems can handle it, store it, and we can use it to get back to the customers reliably. Hopefully that is an interim process, and in the long term we won't be doing anything but taking in raw data, utilizing the raw data, transformation issues that come.

When those situations do arise though, it's important for registrars and registries to have processes in place where manual intervention by support staff can bridge the technical gap that we're in right now. So having an alternate mechanism for validating customers, etc. It's not an uber technical point here, but in other words, have a fall back where there is a support system to bolster this in the mean time.

DON HALLANDER: All right, thanks Rich. So the next section is, how can geeks help? So assuming the room is mostly full of geeks, this is what we think would be useful. Support the Unicode in the local part of an email address, and then the other information that it might be held, so those people can use Unicode in their name or address fields.

Support IDNs at top and lower levels of the domain name. Recognize that there are many, many more TLDs in existence. So one of the issues that we're finding as we do the evaluation of the world outside is, a lot of applications still take the view that a domain name, a top level domain is two or three or four or six or seven characters, when there are many others, and that all of the characters are ASCII characters.

So go back and check your systems to make sure that your validation rules, your filtering, on acceptance are realistic and modern. And then the other thing that would be very useful, because the world is a very big world, there are something on the order of 38 million software developers around the world, is raising awareness within your own software community at home.

And looking at slicing and dicing the world geographically, by industry, by language, by preferred tool set. So, let me talk a little bit about the universal acceptance steering group. Hopefully you've got some idea about what universal acceptance is. So universal acceptance steering group got created just about a year ago. So universal acceptance has been an issue for 15 years, accelerated in 2010, really accelerated 2013.

And the beginning of last year, there were… ICANN has been having universal acceptance sessions at the ICANN meetings for some years, and nothing seemed to be happening. So a group of people said, got together and said, can we make this happen? And a gathering happened in Singapore. And so we now have this universal acceptance steering group, which is a community group that is gratefully supported by ICANN, but it is not an ICANN group.

And the group's primary objective is to help the software development, software developers and the website owners to get their act together. So it's a community group supported by ICANN, not an ICANN group. Active participants include Aflias, Apple, Asia, the dot Asia, ccTLDs, Doughnuts, Go Daddy, and others.

So what we're doing is we're developing documentation, and an approach on how we're going to reach out to the world. So when we did some informal engagement with CIOs and software developers, and they got, they understood the idea fairly quickly, and then they said, "Right. Can you give me some documentation to fix it?" And that's principally what we've been working on for the past three or four years. So where are we going?

We're finishing our core documentation. So we've published this week the quick guides to universal acceptance that give you some good practice guides for acceptance, validation, storing, processing, and displaying. We're about to launch a project to find developers to look at the most popular software development tool sets, languages and toolkits. And see how UA ready those are.

And if they're not, then to fix them. We're looking to raise awareness, and we're facilitating EAI development and deployment. So the next bit is a bit more geeky, and Mark should have been for it and I'll just give my best go for this. So I already talked about the quick guides, and those are available on the UA Wiki page, and you can get to it at UASG dot tech.

And we basically say, accept everything, validate, make sure that you validate for appropriate form. Store the domain names, store the data in UTF-8 format process that way, and make sure that you display in Unicode. We're also working on the UASG 007 document. This is the master, big, ugly, very technical document.

We finished our reading yesterday, so we've finished version seven. We'll make the adjustments over the next week or two weeks, and then we'll publish a version eight and come out to the community and ask people to review, make sure that that's

right. This is the document that you're going to go to your system's admin, system's architect, developers, programmers, to use to figure out how to get your own systems UA ready.

So [inaudible] EAI email address internationalization just should be able to support IDNs, and so when it goes to the wire, it will transform to Punycode, so it can look up the MX records of… Eberhard asked about that earlier today during Marvin's presentation, using Unicode outside the wire, accepting IDNs and displaying IDNs properly, and supporting Unicode in the local part, and that's the hard bit.

And the other hard bit is, if you have Unicode in the local part, what if the MTA at the other end, doesn't have Unicode in the local part? How do you deal with that? And you saw that Core Mail's approach is they're having an ASCII alias to their Chinese dot, Chinese at Chinese dot Chinese. That's one approach. There is other approaches including just say no, you can't, I won't accept that.

So that's a hard bit. There is progress underway, some of the local players are in production in China, Taiwan, Thailand, Saudi Arabia, in the US; and the big players, Apple, Google, Microsoft, Yahoo, they are all actively pursing EAI to one extent or another.

So Google's Gmail is able to send and receive EAI addresses, but at this point, can't support an IDN address in the Gmail as a

Gmail account. I'm not quite sure of the right words. And Microsoft expects to have their products functioning. They're not going to say officially, but Mark is not here so I'll say that by the end of this year, Office 365 should be EAI ready, Outlook 2016 should be EAI ready.

And depending on how you use mail, so some people work straight from the browser, other people like a client. So both should be available, and you heard this morning from Marvin, that they have client software that runs on Windows, Android, and iOS, but not Mac X and not [inaudible]. So stuff is happening, in terms of EAI, because when you talk to your CIOs or whoever is driving this, their basically not really interested until their email supports it.

RICHARD MERDINGER: And something to note, this slide takes me specifically, because I look at it and it looks to me… You know, you think about it from a CIO perspective and the implementation of your own systems, and you're adding your business value to, you know, they're adding business value to products, etc. And the CAI thing, well how many…? Invariably, the question, how many of these email addresses are attempting to get to us and fail?

How big of a problem is it? The big players: Apple, Google, Microsoft, Yahoo, starting to accept these and actually transport

these properly and effectively, means that those floodgates are starting, they're getting ready to crack.

So people should be getting ahead of this, if they want their customers to be able to take the email address that they've associated with their identity and use it for their products, and use it for logging in at their banks, and using it for the full use of an email, like we're used to today, this is the harbinger of that about to become reality. And the pain will be there even, or it can be eased by trying to get ahead of it.

So I don't mean to be marketing about it, but we're at the forefront of this and there is still some time, but the time is now.

DON HALLANDER: So here is some, just some examples of some of the hard parts. Normalization of characters is an issue. Variance that we've been hearing about within Chinese and Arabic, and other languages for many years, are still a challenge. Dealing with systems that are not EAI ready is a challenge. The left to right scripts, Hebrew and Arabic, and shifting from browser to client based solutions.

So what is the UASG doing to accelerate solutions? I said before that we're looking at the top development platforms. This is like Ruby, and within Ruby, the Ruby Unrails, or PHP, or C++, or

whatever development tool you might use. And we're looking at the top five to 10 platforms, and we're going to make sure that they are UA ready, and if they're not, make them UA ready, provide that input where we can.

And the goal here is to make sure that we eliminate any excuses as we can, so that people say, don't go, ah well, I use whatever, I used Ruby Unrails, and it's not ready and I'm not going to develop my own utility for that. So we're going to say, yup, we get that, so we'll do that once for that whole community.

We're preparing and producing and publishing some case studies. We're measuring UA readiness and we're going to the top 500, top 1,000 websites of the world, and we're trying to register, or log in, or create some engagement with a variety of email addresses in different configurations. And so far we have found just one website that is, that accepts all of our email addresses. And just so that you can be a little befuddled, so far that is My Space dot COM. And I think everybody here is probably too old to know what that might mean.

I had to ask children. And we're encouraging local advocacy, and very keen for this community which are geeks in the nicest possible way of the word, to go and talk to your colleagues. And this is, for those of you who are in a Latin based or even English

based country, it is not, it is your issue. It's not your issue. So I live in New Zealand. We have three official languages there.

We have English, we have New Zealand sign language, and we have Maori. And Maori uses the ASCII character set plus the vowels with a macron. So I have a bookstore in Wellington, and so in Maori the word for book is [Maori]. And the word for books is [Maori].

So N-G-A with a bar over it. So I have a domain name called [Maori] dot NZ. And I cannot use my email address which is [Maori] at [Maori] dot NZ hardly anywhere. And until very recently, Internet NZ which runs the dot NZ namespace, I could not become a member with that email address.

I raised it to their attention and they fixed it quite smartly. But so it's an issue for us in New Zealand. For those of you in Europe, or where there is a French community, you cannot have café dot whatever, that's an issue. They don't work, so it's an issue for everyone. And I'm done with time?

So I have one more slide. So this is a summary of where we're at. We've defined universal acceptance. It's a software issue. It's not a DNS issue. It's effortful, but it can be done. The Internet industry is working on their own solutions which talked about the stuff that VeriSign and Go Daddy are doing.

That work should be finished in the next couple of months. And then we're going to start doing a concentric spiraling out circle of including more and more players in the community. Documentation is nearing completion, so the quick guides are there. EAI is the hard part. We ask that you get your own systems in order.

And if you want to subscribe to what's happening, there is a tiny URL dot come slash UA discuss, or if you want to see the documents you can go to UASG dot tech. Just one little warning about UASG dot tech, when you type UASG it may very well change to usage, which is not the right domain.

So that's what… I'll be happy to answer any questions, or happy to move on. Thank you Eberhard.

EBERHARD LISEE: Okay. Thank you very much. A little applause is always in order. [APPLAUSE]

Any questions? Even so shortly after lunch?

JAY DAILY: Thank you. I've got 14 questions. Okay. Jay Daily from dot NZ. First one, you said that the gTLD who doesn't support UTF-8, and

now as you probably know, from many of us in the ccTLDs, just have a UTF-8 WHOIS, which is, we just did it and it just works.

And you can do that looking at the domain name, that was done there. You can look at the dot JP one, just who is just do it straight. Have you had the conversation with the president of the global domains division? Or anyone else senior there?

RICHARD MERDINGER: What is happening is there are discussions going on regarding the evolution of the WHOIS system and until they can batten down the types of details regarding what to do with internationalized contact information, etc. we have to fall back in the standard that's currently in place.

JAY DAILY: Okay. I think you could aim for an interim thing which is just put UTF-8 in WHOIS. I mean, it's as simple as that.

RICHARD MERDINGER: It's as simple as that, and we can take this off. There are operational reasons why the introduction, especially if you look at a, and I'm not advocating for going think thick, vice versa, that's not what we're here for. But if you look at a registrar based community with a thin WHOIS model, and you get part of

them starting to support UTF-8 and others not in the same manner, you have automated systems that support transfers, things of that nature, that could really come into some level of uncertainty.

So until we know where we're going…

JAY DAILY: Okay, I would be in favor of seeing that uncertainty in flushing people out, but okay. When you talk to the… I'm going to carry on.

EBERHARD LISEE: One more question.

JAY DAILY: All right, three more. Have you spoken to the large email providers to get statistics from them on the number of incoming EAI address messages that they can't process?

DON HALLENDER: It is Mark. So we're certainly engaged with the largest email providers, and we haven't asked that question yet.

JAY DAILY: Okay, because that would be useful so that at least we know how many people are attempting to use it, and…

DON HALLENDER: So a lot of this past year has been about getting them engaged and willing to talk to one another. And we're still not there, but moving there, and we can ask the question.

EBERHARD LISEE: All right, thank you very much. Any other questions?

RICHARD MERDINGER: Eberhard, I have a question, if I could. It would be extremely quick. Just a show of hands, how many people in this room were comfortable with the concept of universal acceptance as we perceive it from here, or is this a new topic for you? So if you were comfortable with it beforehand, just raise your hand. Thank you very much.

That's telling for us to help us understand the scope of what we have to do. Thank you.

EBERHARD LISEE: Okay. Thank you very much. Next is Dr. Ricardo Scmidt from the University of Twente.

Ricardo will speak about self-organizing any cast.

RICARDO SCHMIDT: So, good afternoon. My name is Ricardo. I work at University of Twente as a [inaudible] researcher. My talk will be slightly different from what you have seen so far today. It will be really research oriented, as I come from university. I'm here also with a colleague of mine who is in the audience, Walter, which works with me.

And I'm going to talk about the SAND project, which is a collaboration between the University of Twente [inaudible] SIDN labs with [inaudible], and L-Net Labs with Ben [inaudible] and [inaudible]. And SAND stands for self-organizing or self-managing any cast networks for the DNS.

And as the title says…

Okay. And as the title says, it's all about bringing autonomous management into DNS, and more specifically into any cast DNS. The idea is to do that with feedback loop that has been proposed by IDN. And it's called MAPE, consisting of four major phases: monitoring, analyzing, planning, executing, which are quite self-explanatory.

And a knowledge base, which we should feed in some information, and that's where we are focusing our research. For

instance, how to get this information about any cast infrastructure that we want to manage. And that's basically about monitoring, measuring, performance, reachability, resilience. We had, earlier this morning, a talk from our colleagues from Turkey about a DDOS attack.

So we are also interested in managing the any cast structure to make it more resilient to a DDOS attack. So as I said, most of our research focus has been on monitoring, and how can we take advantage of tools that are already available, technologies that are available there. For example, using RIPE Atlas, or any other distributed framework worldwide to measure the any cast infrastructure that we are interested in.

It's interesting to know that we try to generalize as much as possible, any cast. So the focus on the project is DNS. But all the methodologies that we develop, all of the results that we get, we try to bring it also to any other sort of IP any cast. For example, used by CDNs.

So for example, available tools and platforms, as I said, RIPE Atlas framework, we use it for probing any cast infrastructure to understand how it's behaving, reachability to understand how it's behaving in terms of performance. But also, we try to use new technologies, which are actually not so new but they are not widely deployed, at least I don't think they are. For example,

EDNS client sub-networks, the purpose of the origin of the queries that are reaching your name server.

For example, Google already uses it. So we can take advantage of this information to understand also not only the resolvers that are communicating to our name servers, speaking from a name server perspective. But also to know where the origin of the queries that I see.

So I will go first to some examples. I will give some hints on research that we have been doing. I will be very brief about it, so I will be happy to answer any other questions that you have on each one of these topics that I will address.

So this is a piece of research that we, one of the first we have done. This shows about 2,225 RIPE Atlas probes around the globe, and we were measuring reachability of the K-root infrastructure. So if you're familiar with the root servers, K-root is any casted a little bit more than 30 sites around the globe.

And we wanted to understand how the reachability of the any cast infrastructure was, and if K-root was given a good performance to the clients that we're trying to reach anywhere in the world. Of course, there are many other name servers that one can use, but we were specifically looking at K-root.

There is some color code here in this map. Each mark is one of the probes that probed the K-root. Basically green means everything is okay, from accessing K-root from that location. Red means it is still good, but it could be better.

This is specifically in my, in our website, the website address I will give later so that you know that sets now, and you look at slides. I'm also happy to answer questions on the way, if that's possible, if that's okay with the chair.

Also concerning reachability, as I said, Google is using the EDNS zero option, the extension from the DNS, where the query is coming from the resolver and getting into your name server, have the information of the prefix of the region of this query. So we can do like, we can make like a two level mapping of where this queries are coming from to understand the dynamics behind the DNS, and to better plan the extension of your DNS infrastructure, of your any cast infrastructure.

This is one example. So here we are monitoring the traffic, reaching name servers at Serf net in the Netherlands. It's the name server for several top level domains. And in this case, we are monitoring all of the Google traffic from the public DNS, Google public DNS that's coming via the data center in Atlanta. And each one of the countries in this map, basically all of them, have a little circle there.

The blue circles mean that most of the queries came from those countries, but each red circle means that that country also originated queries that reach our name server. But somehow queries from Australia, or from India, were served by the data center in Atlanta, while Google has many other data centers worldwide.

For example, they have one data center in Singapore as well. It served most users in the Singapore location, so southern Asia location. But again, there are spots in every single country around the globe. So if the circles was represented the [inaudible] of the data center in Singapore, basically the circle would have to cover the whole map.

So from this data, we understood that yeah, we can plan an extension of an any cast network only looking at the diverse oversight, but if we want to look also at the query origins that are reaching our name server, we can go one level deeper, and that will become a little bit more complicated because I see traffic coming from everywhere.

Another piece of research, again we used RIPE Atlas. In this case, we used seven 500 probes around the globe to test whether, how many sites in any cast network has to have such that it provides you good performance. Good performance is very abstract. We are just comparing two root servers, C and K.

The idea is not to say C is better than K, it is not what we are evaluating. But basically what was interesting to see was that C has eight sites around the globe, or not around the globe but in North America and Europe. And K has around 30, 33 sites all over the globe, and still the performance was very similar between these two roots, which is the continuous lines both red and blue.

Blue is for C and red is for K. But the dash at nine, which represents the optimal, possible. So let's say that we use the 33 instances, 33 sites from the any cast from K-root at its maximum, then the performance could have been much better. If I am to give you some numbers, the median latency between anywhere in the world to a K-root instance root, would go to 30 milliseconds down to 19 milliseconds.

Still in the same study, it was, we also try to understand how many instances we need to get the maximum from our any cast network that we can have. In this case, we used Z-root, as I mentioned eight instances around the globe. And in this graphic, we were basically playing around with instances of Z-root combining them to see how much good the performance would be with different combinations.

The blue line which is basically hidden behind the red line here in this spot, is the optimal that we can get with eight instances of

C root. And we almost reached that optimal only with three of those instances. So basically, the placement is much more important than the number of instances of the network, that was our main takeaway of this study.

Of course, I'm only looking at performance here, because if I'm talking about resilience, then you need many more instances. Once you enter the performance perspective, very few instances are enough. We are also looking at the visualization. So again, just to go over the concept.

Sandy is working with this MAPE loop, so it's monitoring, analyzing, planning and executing. And analyzing the data, the part of analyzing the data is quite important as well, so we are working in finding ways to show the data so that the operators can understand what is going on in their any cast infrastructure.

Here is a very recent work from me and Walter, my colleague. Where we have, we monitored one any cast system from two sources of data. One is at the controls plan and one is at the data plan. So the plot to your left, the more dense one, we have the very central node is the any cast service being announced in this AS number, which is 47 065. That's a BGP test bed that we used for this work.

So it's real data. Real BGP data. And we used the remote route controllers from RIPE to take a look on how the routing is

actually established, and how it actually should work. How you expect it to work. So everything is working fine until we start to run trace roots towards this any casted service at 47 065, and that's the plot to your right.

So the leaves of this plot, so that other part of this plots are the servers where the trace root is coming from. And the destination is the central point. One step from this central, so the few nodes around the center node, these are the any cast instances. So it's where the trace root hit the service, the any cast service.

And we saw some very weird things going on, that were mis-configuration that the routing configurations of the system, and not at the BGP routing tables. So from the BGP routing table, everything was looking fine but once we were looking at the data plane, with stretch roots, we saw that traffic was reaching their instance in the US, and somehow being rerouted to their instance in Brazil.

That was just a mis-configuration but that can happen, right? So that was one example of how the visualization really helped someone to reconfigure their any cast network. So this is an ongoing project, ongoing effort.

So this is an ongoing project, ongoing effort. We have been one year and roughly four months on this project so far, and we have some lessons learned on the way. So fully autonomous as we

were planning at the beginning, prove to be very challenging and that's because there is a lot of manual things involved on setting up any cast network so far.

There is a lot of BGP negotiations and so on. So it's not just wishing an instance somewhere in the world and making it available right away on the spot for transient traffic or something like that. It's a little bit more complicated than that, and you probably you know even better than me.

But semi-autonomous is something very possible. And as I highlighted, our work has been very focused in monitoring and measurement, and measurements really help us to identify hidden mis-configurations, visualize how the any cast is operating. The only drawback, and I called here drawback but not really drawback, is that usually you have to set up your experiments. But your measurements, not experiments, sorry.

Sorry but measurements, they don't need to be experiments, they can be running for long periods of time. And then you're going to have these organizations on a longitudinal period, a longitudinal scale of your service, which will help you to make decisions on long term for any cast infrastructure.

And changing a bit of topic but not so much, but exactly this moment we are focusing on any cast test bed for our, in our project. We are running SAND, and there is another project that

is touching together which is also based on DNS any casts, but focused on security. And we are trying to create this… We have created a test bed, mainly because any cast has a lot to do with BGP.

And BGP is not something that we can simulate in the office, in a computer or we can, but there is a lot of abstraction going on, and then results might not be really applicable, so you want to run things in the wild and see how things react and how things behave. We have resources allocated by surf net and RIPE so far, so we are not asking for resources, but we want to go as global as possible with this test bed.

This is a research test bed, so traffic is very limited to the nodes that we deployed whatever we deployed them. They are re-delegated to pings, eventual pings, trace routes, and mostly DNS requests. So far, we have eight instances in our test bed, some of them are due to deployed but they are confirmed instances.

Maybe some of you know net [inaudible] which was our first help in this test bed setting up in this instance in London. We had one in the Netherlands, working in North America, in Asia. But as more instances that we have, the better, because then we can play around with combination of instances. We can also understand how we can better mitigate the DDOS attacks that are coming from wherever in the world.

Yeah, if you have availability to host one of our nodes, and by hosting nodes I mean it's a virtual machine running, that we will get us access to announce our BGP prefix, and remove the announcement from time to time. That's it, that's what we need. Please come to talk to us if you can help us, on setting up new nodes.

Many on the spots where we don't have like South America, Central America, Africa, Middle East. Well, that was it. I will be answer to answer any research related question. I'm also happy to answer by email. This is the project webpage. And also the email of my colleague, Walter [inaudible]. Apparently you can send mail only to him because mine is not there, for some reason.

Okay, yeah, it will be on the slides, yes. Okay, thank you very much. [APPLAUSE]

EBERHARD LISSE: You got the applause without prompting. Any questions? Andrew?

ANDREW SULLIVAN: Andrew Sullivan. Don't go too far, because I'm going to ask this quickly and then you'll have the mic back. I wondered how you make decisions about the location of the infrastructure, like the

geo location of the infrastructure when you're doing it. And the reason why I ask this is because when we do something similar at [inaudible], and we found actually that the data that we get, quite the available data is mostly garbage.

It has got all kinds of mistakes in it, and we spend a lot of time collecting it. So what have you done in order to try to clean that up?

RICARDO SCMIDT: Are you referring to any specific…?

ANDREW SULLVIAN: Yeah, like for instance, you said, oh well this is coming from Atlanta. How do you know?

RICARDO SCMIDT: Okay. It's clear we're trusting in geo location maps at this point. That is a very initial research. I generate this plot maybe two weeks ago. So we still have to check how much of that data is actually garbage.

ANDREW SULLIVEN: So are you checking the address in the EDNS zero client, okay. So the end user stuff is actually considerably better. You do this

with the infrastructure stuff, don't believe a word of it because it's mostly…

RICARD SCMIDT: We, Walter is responsible actually for the validation. So we are using RIPE Atlas to validate from all over the world to see if we get the exact information.

EBERHARD LISEE: So identify yourself for the remote…

BILL WOODCOCK: Bill Woodcock here. So [inaudible] was telling me about this research some time ago and was saying that the big concern with it was that because you were dependent on ATLAS, it really only gave you a point of view that describes Europe, right? On the one hand, your publishing this as though it's talking about the world, but all of the measurements, essentially, are taken from Europe.

And so the vast majority of the world, you're not even able to see from RIPE Atlas. And he was talking about there being future work that would pare back the European nodes in order to get parody with the other regions so that, you know, you could have something that would be statistically valid. But that that would

pare it back to such a small number, that then it wouldn't be statistically valid.

Can you discuss the problem, and what else you're looking at to try and get around the Atlas problem?

RICARDO SCMIDT: Yes. So John [inaudible] was one of the core authors of the work, of one of the works that I presented. And indeed, RIPE Atlas is quite biased towards Europe. Most of the probes are located in Europe. And from Africa, for example, you have a very narrow view of what's going on. And narrow view, I mean that you have good view from South Africa, but from the rest, central Africa, you have basically nothing.

So what we are going to work now is some other measurement, where we are going to use our own test beds. That's why it's important to have more and more nodes. Where we are going to help from our own test beds, to the outside and see where the answers are coming from. So basically in theory we will be able to probe the whole IPv4 spectrum, for example.

So we can have a complete view of the Internet and catch which ability and so on.

EBERHARD LISEE: Any more questions? Thank you very much.

DANIEL: You saw some…. Daniel [inaudible]. So, in the picture you saw some RIPE probes that had issue. The red dots. Can you explain?

RICARDO SCMIDT: Yeah, that research was… So if you go to the webpage, there is on the publication link, there is a better explanation. But in summary, there are four colors. The green one were probes that reach the closest instance of K-root, RT wise, latency wise. The red ones, they didn't reach the closest one. That's it.

It doesn't mean that it was a high RTT, but it was not the closest one possible. And that there are many reasons it can be routing policies from local instances, for example.

EBERHARD LISEE: Thank you very much. [APPLAUSE]

And now Jacques can take over and organize the roundtable.

JACQUES LATOUR: Hello. I'm Jacques, I'm with dot CA.

So last time we did the planning for tech day, I said it would be interesting to see all of the different technology offering stuff that ccTLDs and providers do with any cast technology. And then I ended up being responsible for this roundtable, panel, or gathering. So what I want to do is, I'll start quickly with showing what dot CA does, and then one by one, come and present and stand in front, and then we'll do Q&A at the end, time permitting.

I think we're already 20 minutes behind. Can we go to 8 PM or no? Yeah? Perfect.

So today we have a little bit of ccTLDs that are going to present their stuff, and a couple of… Oops, what happened?

And then a couple of ccTLDs, DNS operator, any cast provider that are going to show technology. So the intent was to see how virtualization was done inside any cast infrastructure, if people are using container specific things about infrastructure, how big the nodes need to be.

How many nodes do you need to have in your any cast platform? Like there is more and more, there is the looming threat of DDSOS and any cast is the best prevention for ccTLDs. And there is a lot of different offering. So the goal is to go, try to provide a better picture of what's out there with technology that we can use to be more ready to defend ourselves.

That was the scope tonight.

Okay. So quickly, for dot CA, we have our own dot CA any cast cloud that we built on top of our [inaudible], service offering. We have cloud one which is pretty big, substantial with about 12 nodes. And it's a global infrastructure across the world. So we've got nodes in Miami, Hong Kong, and so on.

So I think that's a standard infrastructure for a cloud. Right now we're working on building our second cloud. Right now it has got only three nodes, and can [inaudible] with transit, but that's where we're at. We're trying to make that one bigger.

So for dot CA we have… The architecture guideline for dot CA is to have a minimum of three clouds diverse. So they need to be diverse technologies. They don't, I don't want all to be the same cloud, with the same technique, with the everybody has a node in Hong Kong or a node in all of the same cities, with the same transit. So we need to make sure there is as much diversity as possible.

We have a few any cast nodes. There is reasons for that because it provides great visibility and so on. And then the thing we [inaudible], that's part of our [inaudible] platform also is that we have Canadian, local, any cast nodes.

That means they're dot CA nodes inside high ISPs in Canada, that only have access to the Internet exchange point traffic. There is no global transit on it. So if we're under attack, whoever is spearing a DIX will still have access to dot CA.

That's something we're building internally is to build more peer at the Canadian ISPs. So the node we have dedicated hardware for dot CA. We have 10 gig router with transit to the Internet. The same transit provider across the world for the global node.

And then we also have a 10 gig peering at ISPs cross, for each of the nodes. That's been the way of dedicated stat collector for each site. But that's the architecture that we have for dot CA. And we keep, our goal is to make it better and more resilient. That's it. That's it for dot CA.

So Bill is going to do a quick presentation, roughly five, 10 minute on basic 101 on any cast, and then he'll go through his PCH specific infrastructure.

BILL WOODCOCK: This is a really old slide deck I'm just going to try and run through really quickly, talking about the interesting issues when you use any cast for DNS. I'll just kind of look through a bunch of these because they're already, you know, well understood.

So the two big benefits of any cast are failover redundancy and latency reduction. And the tricky thing is that that interacts with the failover mechanism in the resolvers themselves. And so you want to make sure that you're sort of getting the best of both worlds rather than the worst of both worlds.

So, the resolvers in the clients are of course working with a list of authoritative servers, and they'll failover from one to the next. Whereas any cast is using the same IP address and is using the routing system to get to the nearest instance, regardless of whether that instance is up, or the highest performing instance, or the least loaded instance, or the one with the most up to date information or whatever.

So you're trying to get these two to work together well. The resolvers are measuring by latency, any cast is by hop count, and those are not going to give the same answer every time. Any cast is always going to trump what the resolver does, if you allow it to. So, what that looks like, let's say we've got a satellite connection that's one hop away, but very slow, and a fiber connection that's three hops away and very fast.

If you've got two servers that have the same routing policy, any cast is always going to chose the short hop count, high latency, low performing connection. The resolver would choose the low

latency, high performance connection, even though it's a greater hop count, because it's not looking at hop count.

It doesn't know about that. The problem is that if the two servers have the same routing policy, any cast is always going to trump the resolver. So the way to solve that is to use two different clouds, two different IP addresses that go different places, like this. You split those IP addresses into a cloud A and a cloud B, and this allows the resolver and the any cast to work together to get the result that you want.

So what that looks like, sorry I'm going… I'm trying to go very fast because these slides have all been up on the web for 15 years. So. Sorry, there is another issue that you want to be really aware of when you're doing a deployment, that also similarly effects how you build out the any cast nodes, and that is that the ISPs will always deliver traffic to their customer, in preference to their peer, and will always deliver to their peer in preference to their provider, right?

So the Internet Service Provider, any time they have a packet to deliver, they're trying to get paid for that. If they can't get paid, they're trying to deliver it for free. If they can't deliver for free, then they have to pay to deliver it. Right? So that's always their order of preference. And that, again, trumps the topology.

Right? The order of payment is going to have a greater effect on routing then the hop count.

So again, you've got something that's not necessarily visible to you that is going to affect where the packets go. So let's look at how that plays out. Here we've got two transit providers, one in red and one in green. You've got two any cast instances, which are both connected to the green provider, all right?

So you're buying transit from the green provider for each of these any cast instances. The red provider has a customer who wants to send you a packet, and the sort of straightforward way of getting the packet to the any cast instance, is the red provider is going to talk to the green provider through the nearest exchange point, and the green provider is going to drop it into your any cast instance, because the green provider wants to get paid.

They're not going to give it to you over their peering connection. And you are maybe not peered directly with the red provider. Okay. So the problem comes if you go and you buy transit from the red provider at a distant location. Now they're always going to deliver over that transit connection, at the distant location, in preference to the nearer location, right?

So the red customer in the east is going to get served out of the west any cast instance instead of the east any cast instance,

because you went and bought transit from the red customer's provider there. So if anybody has questions about that or it wasn't clear enough, download the slides or I'll take it through later when you have more time.

So, peering with red is not going to fix this problem because as long as you have a transit connection to red, they'll never use the peering connections. So you have to get the same transit providers at every global location in order to fix this problem, this is, this makes transit provider selection for an any cast network really difficult because you can't grow the footprint of the network any bigger than footprint of the intersection of your providers.

It also means that you need to choose the providers really carefully so that they're all large, and they all have a large degree of overlap between the places where they can provide service. This is especially rough in Latin America and Africa, where there are not so many providers that have good global footprints, who also serve those regions.

Okay. So how does all of this fit together? We need redundant hidden masters. Each of those is serving their own cloud, so an A cloud and a B cloud. We use something called dual wagon wheel topology, we call it that, where each node knows about a couple of other nodes in that cloud.

And obviously the hidden masters have to be synchronized with each other. We have to have redundant transit, so we use transit that can hit all of those locations, and the other alternative is to use one pair of transit provider, providers for one cloud and a different pair of transit providers for the other cloud.

This is one way of getting around that problem, although it gets to be a pretty big hassle to manage, because now you're looking at four providers rather than two, and if you have provider, you know, A sub one, B sub one, sorry. A sub one, A sub two, and B sub one, and B sub two, you can't mix and match A sub one and B sub two into the same cloud. Right?

You've got to stick with the same pairing throughout. So then we have local peering at all of these locations. So how does this work with resolver based failover? First the resolvers, the customer resolvers, are going to pick whatever gives them the lowest latency, and that's the desired outcome.

Then if a node goes away, they're going to failover to the next quickest to answer. So again, that's the result you want. Any cast, internal routing is going to give you the failover to the next, nearest node, which is again, what you want. This is the advantage of that wagon wheel topology.

And then from a global, an external any cast perspective, once you've been able to withdraw the routes that were supporting

that node that went away, then the customer will wind up going directly to the next top count destination. So, when there is an attack on the any cast network, obviously all of the server nodes are reachable by the attacker, and legitimate users get blocked. When you've got any cast, you know, to some degree you get this beneficial effect of the attackers hitting some nodes and impacting legitimate use in those nodes, but having other nodes that are not affected or not affected enough to be noticeable, and having no detrimental effect on those users.

Okay. So and I'm going to move on to talking about our sort of specific service offering and so forth. And again, questions for the panel or afterwards or whatever, on the general stuff. So PCH is a non-profit, been around for a bit over 20 years.

We're industry supported to deal with Internet infrastructure issues, critical infrastructure issues. We have been any casting TLDs since '97. We've been providing any cast infrastructure to root operators since 2001. And we've got a DNSSEC signing platform for ccTLD operators that we set up with ICANN about five years ago, that is [phips] 140 dash two level four, like the root.

We are currently are providing infrastructure for three of the root operators, 532 TLDs of which 103 are ccs, the rest are gTLDs including 122 IDN TLDs. Overall, 7200 zones hosted. Most of the

ones that aren't TLDs are conical SLDs. So like, you know, co dot foo, and org dot foo, and edu dot foo, that kind of thing.

Or critical infrastructure, so exchange point domains, certs, governmental domains, and IPv6 ARPA, and [inaudible] ARPA hierarchies. So in terms of our topology, what we look for from registry is at least one hidden master, ideally two or more, where those servers are visible to us but not to the public, so that we can't lose our source of authoritative zone data due to them being DDOS.

Then we have a set of intake slaves that receive updates from the registries hidden masters and push those to our outbound masters, and some of our internal monitoring systems and so forth. And then we push that out to our any cast nodes. In addition to that, we've got our DNSSEC infrastructure, so we take unsigned zones in and push signed zones out. And then we've got a measurement infrastructure, that also receives IXFRs from the registries hidden masters, and then starts a timer on every update that we receive, and starts watching the, each any cast node to receive that update.

So we get an update from the registry's hidden master, and then we start watching for updates back from our own any cast nodes, and if we're not seeing those, we start polling to try and figure out what the problem is, or why it's taking too long, and

obviously to maintain statistical tables about the propagation times.

130, mark 135 locations now. So how we divide that up size wise. About 70% or what we think of as small locations, that handle up to about 250 megabytes of traffic. The next 20% are medium locations that handle 20 to 60 gigabytes of traffic. And at any given time, about 10%, so right now 13 locations, are large, handling 60 to 120 gigabytes of traffic.

We pre-configure every location before it goes out, all of the equipment. We've got sort of the usual operational code that manages deployment stuff. The small locations are all host self-installed, right? So we're putting together a hardware package that's preconfigured, shipping it out with single page of instruction, which is just a little cartoon showing where to plug in the four cables, three Ethernets and the power, and you're done.

So very much like what you're hearing about from F-root, from Brian this morning. And the medium and large, we send our own people out to do those installations and upgrades and so forth. Small, right now is the Cisco 29 21 router, with a UCSE 160 blade server inside it. So this is a X 86 server with 64 gigs RAM and two one terabyte SATA drives, that fits into the 29 21 router.

So it's a single integrated package. Uses the same power supplies and power cable, all internal management on the back plane. The two devices don't have to be cabled to each other, right? They're in the same box. All of the connections are internal on the back plane.

So that gets plugged into typically two gigs of peering and one gig of transit, but again, the router itself limits that to about 250 megabytes of actual throughput. The medium installs are Cisco ASR 9001 router, a pair of UCSC servers, which are 768 gigs of RAM and eight one terabyte SAS drives, and a Nexus 34 48 switch, which is a 48 port 10 gig switch to glue things together, and handle, you know, any additional cross connects beyond what can be plugged directly into the SAR 9000 and so forth.

And this setup we install if we've got sort of a minimum of 20 gigs of connectivity, and it can handle a maximum of about 60 gigs gracefully. And then the largest 10% are SAR 9,006 router, three to eight of those same UCS servers, depending how big of a location. 93, 96 switch which is, it looks a lot like the 48 port 10 gig switch with then a 12 port 40 gig switch stacked on top of it.

So we can do 80 and 120 gig interconnections internally by lag grouping the 40 gig ports together. And so we do this when we've got say 40 gigs of bandwidth that we need to service, up to about 120. Software wise, this is all VMWare ESX right now, and

has been ever since we went off a blade servers, moving towards KVM, run [inaudible] DOS, bind NSD is the name servers, and all of the management monitoring is in house code.

You know, we have used parts and pieces of various open source systems, but there is nothing that, you know, fits this exactly. So. So a few of the things that make this interesting or different. We were essentially making our own bandwidth, in the same way that, you know, a Dutch telecom, or an AT&T, or NTT is. Internet bandwidth comes from exchange points, and transit is the service of moving that bandwidth from an exchange point to where it's consumed. And so that transit, that service, is making it more expensive, it's making it slower, it's making it less reliable.

So we only install into exchange points, where we do not have to use transit in order to keep those costs down, in order to keep the performance and reliability up. One of the other advantages to only going into exchange points, is that we don't disrupt local markets. So one of the things that you often see that is problematic with CDN deployments is that, you know, they'll have to sort of meet their targets for another location that they want to deploy at.

And they'll find that the easiest way to deploy there is to go to the incumbent and say, you know, drop us into your data center

and plug us into a transit connection. The problem is that that incumbent is probably already market dominant, and if you give them yet another data resource, that sits behind their AS in their market, then this is one more thing that reinforces their market dominance.

It's one more thing that they can hold over the heads of all of their peers, and all of their potential peers, and new market entrants and say, "Well, we're not going to peer with you. If you want to get to that root server, you just have to buy transit from us." And that makes the Internet market not work well. It is really harmful to people who want to use the Internet and have to buy the service.

So as a matter of policy, we only put content up in such a way that it is equally available to all comers. That no one has to purchase transit to get to us. And that's worked well. We have about 8,000 direct connections with other networks in 130 locations on six continents, and we're adding more all of the time.

One of the other things that… This always seemed to me like kind of a brain dead, like everybody ought to do it this way, I'm not sure why people don't, things that people seem to like, that we do is auto configuration of new zones. So when we have someone who is using our service, and they want to add a new

zone, and have us any cast it, they configure it on their side, which means that their side starts sending us updates.

When we see an update for something that we have not seen before, we auto configure it across our system, and start serving it. And if that zone then goes stale, we check to see whether it is delegated to us down from the root, if it not, we pull it out.

It means that nobody has to write to an API, or send us a spreadsheet, or email us to tell us that they want something done, you know, they just do the part that they would have to do anyway, and we pick it up from there.

Internally, we convert everything to IXFR, if it started out as AXFR, just makes it more efficient, particularly in places that have little connectivity. And that's it. Sorry, I ran over time.

JACQUES LATOUR: Patrik.

PATRIK FÄLSTRÖM: There. So Patrik Fälström from [nat] node. We also provide DNS services over any cast, otherwise, I guess, I would not be here. Thank you Jacques.

I'm not going to say much about our infrastructure because it's very similar to, I would say that our infrastructure is very similar to the one PCH is providing. We also provide our nodes.

We also provide our services at our access in a similar way. We are also providing service, we are providing services also divided between smaller and medium sites, everything from physical sites running on the silicon itself to smaller sites, also, using virtualized systems. So not much different from what PCH is doing.

I could like reuse your slides to argue in favor of what we are doing. But what I was going to talk about was something else. We have about 60 sites around the world, and we are like PCH peering, we have quite a lot of [inaudible] all over the world, including many [inaudible] in multiple locations.

And I wanted to show, we rely on [inaudible] and DNS and ANTP and other kind of stuff. So anyways, what we are doing, blah, blah, blah. So what we are doing is that we have been, I've been looking at the data that we have collected for the day in the lifetime, because we also run I-root. And looked at where queries are coming from on the different locations according to both geographical and natural topology.

And I think what I'm going to show here is sort of similar to the risks that Bill was talking about, that if you sort of don't control

what you're doing, both if you're the any cast provider, but also if you're running an ISP, if you don't think about your exit strategy, it might be the case that when you're peering with an any cast provider, in that case weird things might happen.

So I will show you data that is sort of proof of what Bill just showed. So, what we have been doing, we've done it now twice, is that we've been looking at the source for an IPv6 address, and we're looking at the… In this case, I was looking at just the 10 [inaudible] we got traffic at, at each site.

I will present what economy the source traffic is coming from based on the RIR information, and it's based, this time it's based on a snapshot in April 2015, which means that it is a bit old, but the next snapshot made now in the spring. So of course, it's not 100% correct, but it's some kind of representation anyways.

So one could say that most of the traffic is okay. We did this also a couple of years ago in 2012. So we now have two measurements done exactly the same way that we can compare between. And it looks better 2015 than 2012, but it's still certain things, some things that one could be a little bit worried about. There is 5.6% IPv6.

We see 1.5% will receive 1918 addresses, and that of course, should not be there. We have the highest numbers in Mumbai and the lowest in Paris. To some degree one can say that as

we're running a root server, if people are caching things correctly, the question is whether we should see any traffic at all, a part from the junk, but that's a different discussion.

So some IXs though, where we have, where we are looking at the traffic, they're coming from places far, far away, and I will show you a little bit of that, a little bit of those examples. Nodes with a lot of traffic, a traffic for more diverse locations, so obviously people choose an exit strategy that prefers large locations, that might well be what Bill talks about, that large locations are things with also low latency, or that is what people think, so they sort of prefer the big IXs in the world, just in case because of course, all of this is, a lot of this is manual labor.

If you have a smaller IX on the other hand, you have a much higher percentage of local traffic, because the local IX gets sort of, the local IX gets local preference, or get high preference from ISPs that are very local, and know that it exists.

And this is also one of the sort of issues that IXs have, that if you have a small IX you don't get so much traffic, but if you just raise above a certain threshold, suddenly you attract a lot of traffic and you grow faster. And we see ourselves, that among our IXs that we have one now in Scandinavia between Sweden and Denmark, which is just now on the threshold of becoming one

which actually attracts traffic by itself, which is kind of interesting.

And then of course, we have also some global eyeball providers, and they are clearly [inaudible] at some IXs, so let's have a look at what it looks like. Here is some examples. And from, it looks sort of, pretty okay. We have [inaudible] present other things, IX, this is one of the largest sort of exchange points in the world.

Traffic from all over the place. Kazakhstan, okay. Maybe Amsterdam is closest to Kazakhstan, who knows? The traffic from the U.S. to Amsterdam is coming from Microsoft. That's basically the only one running traffic there. If you look at Ashburn, another one of those really large places, which, where there is lots and lots of traffic, most of the traffic comes from the U.S., of course.

If you look at the different ISPs, Google by far the largest, but of course, we also have telecom Argentina, that think that if we are going to send traffic to I-root, let's send it to Ashburn. Okay. Yeah.

I can tell you, it is not that, we do have other I-root instances than in Ashburn that is closer to Argentina. Brussels, Taiwan, okay. Taiwan is not very close to Brussels, specifically as we have an I-root instance in Tapti, but ISPs in Taiwan send I-root traffic to Brussels.

That of course means that [inaudible] on [inaudible] time in Tapti, in Taiwan, looks a little bit weird when they look at what, where we are. This is the [inaudible] IX. Small one, specifically last year, it's much larger now, but it's still very small.

And this is one example you see, traffic is local because no one else sends traffic there, that wants direction over the IX exist, and they are the ones which RSPs which really are there. So even though there are global [ASS?] that connect at [inaudible] and other places, their exit strategy don't send traffic there unless it is something, a prefix that is unique there.

So this looks good. Another example, anchor. No one else then, ISPs in Turkey send traffic to anchor. Beijing, most is China of course. Tokyo, on the other hand, whoops. And you see 90% of the traffic to Tokyo is coming from China, even though we have a node in Beijing.

Yup.

UNKNOWN SPEAKER: Maybe the content in the Tokyo node is more trust worthy?

PATRIK FÄLSTRÖM: Yeah, they talk about how trustworthy the traffic is. Maybe it is, who knows? Remember though that we are, I am here looking

at the packets that actually reaches our statistics collector box, and also mention that I didn't say anything else then the traffic that we've chosen to collection box, okay?

Given this traffic routing thing, I have a separate presentation where I can talk about what traffic reaches the IX in Stockholm, and who exchanges traffic there. Not only ISPs inside Sweden, people abroad change traffic in Sweden then in some other countries close to Sweden.

Pairs, of course most traffic from France, looks good. And also if you look at the various ISPs, it's pretty good spread. The [inaudible] is the largest provider. And then if we look at Oslo, 88% local, that's fine. But if we look at IPv6, we see that Oslo, for some reason, I can go back again.

88% of the traffic to Oslo over IPv4 goes there, but 25% of the IP6 traffic is not coming from Norway. And if we look at the various, let's see if I have… Yeah. So for example, there is IPv6 traffic for the [inaudible] ISPs in Thailand, they send their traffic to Oslo to exchange traffic with us.

And the same thing in China. So for IPv6, Oslo is the new black, just so you know that, all of you. So, anyways, so my conclusion here of, I can leave this one here. So, my conclusion is that as a follow-up to what Bill just described, parties that have large aliases, which peer with multiple aliases, just like Bill was talking

about, really should look at the exit strategy that they are using, when they are applying the various prefixes that they are injecting in their IGP, because large [ASS] as far as I see, too many of them have just, they're just a single exit strategy, specifically for IPv6.

And this is also the big difference we see between 2012 and 2015, is that IPv4 was [inaudible] in 2012, but definitely IPv6 was cracked in 2012, IPv6 is sort of nicer in 2015 but still pretty bad. So that is something to think about even more. Thank you. [APPLAUSE]

JACQUES LATOUR: Thank you. Next up, Andrew Sullivan. Is the web X back on?

So for people on remote, we're having a technical issue with the Adobe Connect.

ANDREW SULLIVAN: So I guess I can start talking while we're waiting for this. Basically what I want to talk about today is… So first of all, I guess you maybe don't all know, I'm Andrew Sullivan and I work for a company called DYN, which is spelled D-Y-N because I don't know, Yankees can't spell.

And I currently live in the US but I'm actually Canadian. And I'm going to talk a little bit about what DYN does, the way we do things. But because this is this crowd, I wanted to talk a little bit about how it relates to TLDs, but particularly the problems that we saw, before that we see all of the time at TLD services.

So I'm going to talk a little bit about how the approach that we're taking can be used to defend your infrastructure, and you could do this to. This isn't that hard. So what I will describe will be something that I think, you know, other people could do if they wanted.

Do you need me to put this up instead?

The problem is of course, it's a Google presentation and it doesn't really want to be nice.

I'm glad that we run the Internet around here. So rather than bore you with all of the slides, I'm just going to talk a little bit, and if this ever comes back, that's fine.

The key thing here is that we have, we already heard about how any cast is good and it solves all kinds of problems. And, oh look, now it's up. If you want to, you can go to the next slide. If you want to run a big infrastructure, I think that the fact of the matter is that you need to do it in any cast now.

There is not really another option. You know, there are all of these good reasons. Next please. But the thing is, it's not magic. And we've seen this on more than one occasion. So you can still do collateral damage to an any cast system.

Right? It's not like any cast is some sort of magic thing. So we've got mister [yuck?] down here next please. And you can tell that mister [yuck], I don't know how many of you remember mister [yuck], but mister [yuck] was this thing in the US. I mean, I'm not from the US, but we used to get it. And it's the bad guy. So the bad guy is agreeing with the nasty eyes.

And so mister [yuck] wants to attack this stuff over here, and mister [yuck] towards this service one, and you've got this on common any cast infrastructure. And because of that though, because it's a common infrastructure, what happens there is your user there becomes unhappy, because they're trying to get a service on three.

So you could do some work with service isolation, but it isn't kind of magic thing unless you have adequate modularity. I'm going to skip this over this very quickly, and I'm going to show you what we do. So basically what we do is we've got a bunch of slices on a system, we call them slices, but basically they're just container instances that run on particular pieces of hardware.

And what you do is you basically advertise these slices in the cloud. So in this case, there is only one slice, but it's advertised as just part of this service group. And so there is this service interface to the outside world, and that service interface is what everybody in the world sees, but you've got this slice.

And so if you want to add multiple services using this same infrastructure, you put them in different slices. And the different slices have different interfaces, in fact, they have different networks. And the idea here, of course, is that you can separate them, so you can announce them differently.

And this allows you an enormous amount of flexibility in how you move these things around, because if you don't like this arrangement, you can just add other stuff. You can add other services on the same, in the same environment. You can have the same service running in multiple slices.

In this case, there are these two black bands, there are two slices. So there are all different things, it can all be running in different containers in the same physical infrastructure. But of course, the obvious thing to do is to start adding more infrastructure to the same arrangement.

So you add multiple machines to this kind of thing, and you've got these services that are running across multiple machines, each one in a separate container. So these are just plain old

docker containers, or whatever container technology you like, and you can move those around however you want.

Now what you've got are these little pieces of Lego that snap together, I'm sorry about the Lego TM, right? But you can snap them together in various combinations. And the reason that you want to do that, is because you can then get a budget for the whole system.

So what you do is you say, I am running, you know, TLD dot TLD, and I want to have this much traffic for this service, and I expect that the service is going to be this big, on the platform. So the zone is going to get so large, and I'm going to have so much traffic with respect to the changes and all the rest of it.

You give it a budget, and you say you've got to keep within this budget. And the system then can manage itself. We heard about this a little bit earlier with the, with the efforts on self-managing any cast. Well this is basically what it looks like. You have a little engine there that notices what the budget is, and it fixes itself.

The other nice thing that… The other thing that's good about this is because you have multiple containers, if one of them goes away, it's not that big a deal. You just, you know, pick it up elsewhere. But more importantly, you can add containers really easily. So when things start to go wrong, and this was the

experience that we had with, in support of TR earlier, you heard the presentation this morning.

When it was time for us to stand that up, it was not a big deal, right? You just pop these things on, because you've got the containers ready to go, and you just add more. You start to see more traffic then you were expecting? Well, turn more up. It's not a big deal.

And these should be disposable. Containers should never be a thing you maintain. You should never log into one. You shouldn't do any sys admin on them. Nobody logs in and does stuff, you just generate it automatically. Automation is your friend here.

When you've got that kind of automation, it doesn't matter. You want 20 more? Add 20 more. So what we found is that you can start really small with this kind of arrangement. You can just stand up a few of them, and it's not a big deal. And then you've got different kinds of options.

So if you have demand and it's expanding around the world, which is the kind of thing that happens for us with, you know, enterprise customers or whatever, then you just stand things up a little bit closer to people, and that's better. But if instead you get a sort of flash crowd, like what we saw in the DDOS attacks,

then you stand up a lot near to the sources, near the sources of your problem.

I would put them here, but of course that wasn't where they were. And you sink the traffic. And by sinking the traffic, they're close to the source, it keeps things across the Internet much better. I think that was part of the point that Bill was making about using IXs.

I should say that this strategy is not for everyone. Building this kind of thing yourself, you need to be doing an investment in infrastructure. So if you're doing this for one particular service, it's really kind of a waste for money, because what, you know, the whole point of it is that it's extremely flexible, and it's not that useful if you're doing, you know, single purpose kinds of things.

Also if you don't have the staff that are used to automation and so on, get them used to doing that before they do anything, because otherwise you're going to find yourself in big trouble in the middle of the night. So this is really an automation kind of system that you need to be able to build.

That's pretty much all I wanted to say about this, and I'm going to stick around with the rest of the panel, and we can talk about this later. Thank you. [APPLAUSE]

JACQUES LATOUR: So next speaker, Jaromir, CZ.

JAOMIR TALIR: So hello. My name is Jaomir Talir. I'm from NIC CZ. So I'll give you a quick overview what do we do regarding our DNS any cast. I'll talk about the IP structure, geographical structure, what we do regarding [inaudible] and diversity. We also offer a small DNS hosting. And IDN's future.

So we started to run our any cast in 2009, when we got four any cast prefixes from the right. At the time, there was a policy that the TLD could have only one any cast prefix, but we managed to change so we've got two any cast prefixes for TLD and two for [inaudible], so in total, we've got four any cast addresses for IPv6.

And we named them creatively A, B, C, D. And this is NIC CZ. And until 2013, we ran also one any cast in Vienna. You think, universal AP address, and then we decided to drop this unicast in favor having stronger any cast. But unfortunately this change was blocked by IANA for minimum diversity check.

So what we had to do is that we, we asked RIPE for another AS. And we've got it, and unfortunately there is a requirement that a

new AS should have unique corrupting policy, which is a bit limits the using of this AS.

Because we wanted to have our any cast structure that we could use all prefixes at the same location, so the routing policy is expected to be the same. However, as far as I know, web is not watching, so this is not an issue for us right now.

But you see there is actually an issue that if you want to have your network under the control, and you want to serve DNS in a [inaudible] from your own AS, you either have to fight with IANA or fight with RIPE.

However, there are some planned changes in IANA technical checks to relax this rule a little bit. This is what we're changing in the future. Right now we are serving three hour any cast prefixes from our standard AS, and the one from the other AS.

Geographically, we have some DNS in the Czech Republic, and some broad frame, we have four locations. In Europe, Vienna, Frankfurt, London, and Stockholm, and for our locations outside Europe. We have different mix of presents in some exchange points.

Some [inaudible] in Tokyo, we are only present only locally, in local ISP or exchange point. And usually, the four instances outside of Europe, we selected by some sort of partnership,

either it's ICANN IC or with NIC Chile, for example, that we just decided, we asked them, or we had reached to get mutual together, to get, that we have put our servers in to their infrastructure, and we also put their service into our infrastructure, so that sort of exchange of physical server hosting.

In every location, we have two physical servers, which both announce single any cast prefix, there is no load balancing. One of these BGP connections always announce prefix over metrics, so only one server is active in one moment.

And so we configure it that this BGP that the host expects all of our prefixes, so we can easily switch to another any cast prefix so we can even put all of the prefixes that is being served from one location. If this would be the case, for example, fighting with some BGP DOS attacks.

On our website, there is a map about location of our locations, and prefixes, and mapping with new location, prefixes. We did some sort of measurement of reachability in the past, that was based by selecting something about addresses in all of the countries. We took this IP addresses from our [inaudible] statistics.

And we are trying to ping from our any cast one node to these addresses, and listening on the other instances of the same any

cast, which one will receive the answer, and then we will make sure the response time of this packet. So we had a [inaudible], but we are not using this anymore.

Currently we think that RIPE Atlas is much better to for such an [inaudible]. Now, in Czech Republic, in the local environment, we did a big upgrade at the beginning of this year, or at the end and the beginning of this year, when we migrated from six servers to 15 servers, from one gig traffic to 10 gig ports.

This was Chile based off our [inaudible] tests that they did for a group of 10 gig network cards, and from a lot of them, the network cards from in total selected as the best one, so we are using this card for our DNS infrastructure in Czech Republic.

In the service in Czech Republic, even though they are also geographically distributed between different locations. They are all in single [inaudible], and there are load balanced by BGP multi-pass. Each of these 15 server, again announce single prefix, we have five of As, Bs, and Ds. All of these servers are served by three different [inaudible] providers are offering 10 gig connectivity.

And we are also connected to our local ISPs in Czech, in Prague, into NIC CZ, and in [inaudible] in Slovakia into [inaudible]. We are trying to have quite strong diversity so we are, our DNS

service runs different operating systems [inaudible] DSD. And we use different DNS servers, not NSD and Bind.

And all three actually are installed in this server, so we can easily switch from one to another. And the same is for BGP service in different service abuse for an open BGP, and all of this orchestrated by the [inaudible] from the center point.

Even in the hardware, we are trying to have some diversity. The service are from different vendors. We have DELL, HP, and Intel. At the place where we have control over [inaudible] and switches, but in the Czech Republic, we are using switches from Juniper and Cisco, the same for alters.

We tried block aide, but we are actually not satisfied with the result. And we offer some secondary hosting for some less developed countries. For Angola, Tanzania, and Macedonia. It's free of charge. And mainly because these partner registries are also using our open source registry threat.

So if maybe anyone is interested to participate also in this project, you're welcome. Regarding some thoughts about future, we always think about the proper distribution of prefixes that we want to run more measurements than we are doing to find the best distribution.

You could see that all our service, our physical servers, we are not using virtualization any more, but to have more locations probably this would be the way, how to grow. We are, of course, looking for more locations, maybe in the partnership mode that we can offer some space in our [inaudible] Czech Republic in exchange for the [inaudible] space in some other locations.

We also thought about maybe providing our DNS service or major ISPs in Czech Republic to even get the TLD closer to customers. But this is really just initial thought. Also that's all from me. Thank you for listening. [APPLAUSE]

JACQUES LATOUR: Thank you. Interesting. So next presenter is Nominet, Chris Griffths.

CHRIS GRIFFTHS: So while we're waiting for these slides, I am Chris Griffths, I'm the direction and GM for the new product and business development team, Nominet. So we're building new applications and new services run our any cast and DNS platforms.

Little mood music.

Anyway, I'll get started while we're waiting for that to come up. So just to kind of, I had some of my slides seem to repeat. I guess the continued message of, you know, the need for any cast, I think to echo what Andrew said. I mean, it's obvious at this point, to point any cast DNS servers is not a viable option at this point.

It's, you know, it's a necessary endeavor especially with significant security attacks and other service interruptions that happen on network. You know again, as you're looking to deploy, you know, infrastructure whether it's DNS or other services, obviously looking at global reachability as well as, you know, your design and implementation obviously that makes…

All right, so next slide please. We'll jump right in. A little bit about Nominet, we're an international Internet company focused on delivering public benefit. We run the dot UK registry, which includes several of the sub-domains as well as a few gTLDs. Next slide.

So any cast, I think this has been covered in depth at this point so we'll go to the next slide. I will point out, any cast is not trivial. So while I think you've seen a number of different vary deployments, you know, each of the panelists have talked about it, it is a design and implementation that needs to be thoughtful.

You need to think about global reachability transit versus peering, how those impact, I think, Bill really touched on the impacts of routing specific packets between different networks and the reachability of those packets, depending on where the traffic is coming from.

So again, I think it's important, you know, good network monitoring and you know, thinking about when things go wrong. So we'll go to the next slide, and of course, I always enjoy putting something from *Airplane* in a slide, but you know, again, things are going to go wrong. So we'll talk about DDOS.

So next slide. So here is an actual real live example of dot UK. And this is a DDOS, not a huge one, you know, again, I think it was a 30, 40,000 QPS per second, but you know, it kind of indicates you know, again, the…

Again, this is going to happen. This is going to be, you know, what you are going to see in the real world, you know, again and we see kind of adjusting traffic and volumes all the time. So this is something to be, you know, need to be planned for, need to have the capacity to support it. So next slide.

So where to put all of the traffic. Again, it comes back to the design, thinking about you know, good network monitoring, having scrubbing or other network services, whether it be transit

or if you're cross peer, making sure that you have the ability to take all of that traffic and put it somewhere.

So either black hole it, deal with it in your design and prepare for failure, because it's going to happen. So next slide. So, I think we've touched on the benefits of any cast. I'll point out some real world examples of what maintenance looks like. So you know, we have eight sites right now. We're in the process of turning a few more sites on at run dot UK.

In particular, this was an example back in January when we actually did some maintenance over the course of, you know, a little part of a day as we were doing some changes inside the network. We took offline one of our data centers in Manchester, and basically you can see the traffic drop off.

And if you go to the next slide, basically the volumes increase. The network continues to operate, things are still operating the same IPs were still answering. So again, these are the kind of benefits that you can do. You can take whole data centers offline.

You need to obviously plan for capacity changes in between those data centers, it's obviously a pretty critical thing in between when you're making and swinging traffic on the network. But to something to think about as you're designing this. Next slide.

I think it has been touched on a couple of different times, but multiple vendors equals diversity. I think it's, you know, clearly important to make sure the diversity whether it be in the hardware, transit appearing, and obviously your DNS software. You know, having at least two different vendors and making sure that each of those, you know, because everybody is going to have zero day, and other issues when…

And when that happens, you know, it's basically hair on fire and trying to fix your network. And that's going to happen all of the time. So having diversity and having the ability to keep your network up, when any of you vendors decide to go sideways on you is always a good thing. So next slide.

Just a little bit about our platform, talking about our data centers. We have eight of them right now, a couple in North America, and the rest of them are in the UK and Europe. Basically we look at, you know, how we handle prefix delegation out of each of them. We've got different and very DNS transit, as well as, you know, varied hardware and DNS software per se.

It's, again, we're preaching or actually actioning [sic] on what we actually preach as well. So just making sure that we have diversity across our platform. Next slide.

Also I think PCH and others that have, have and support and distribute via secondary's is obviously a huge link. I mean,

getting and further spreading your network wide so you have a larger service area against attack, and being able to absorb larger spikes in traffic is pretty important.

And making obviously globally available. DNS resolution is obviously, in particular for this, so look at your design. And certainly look at options to push and distribute your data further. So next slide. I think that's it. Thank you. [APPLAUSE]

JACQUES LATOUR: Thank you. Are we…? The last presenter, Olafur from Cloud Flare.

UNKNOWN SPEAKER: While Olafur is getting established, I should state that I have some of these stickers about the IETF 30th anniversary. So if you want one, you've got to come and get one from me.

OLAFUR GUDMUNDSSON: Hi. I am Olafur Gudmundsson. I am from Cloud Flare. We are a content delivering effort first, but we are also one of the larger DNS providers in the world. And so far, we have not provided service to roots or TLDs, but we are open to discussions on that. Next slide please.

Oh, I get the clicker.

Okay. So, right now, we provide all of the services, we provide over any cast, no matter whether they are on UDP or TCP. We believe any cast is the way to go. We realize that operating any cast is difficult, so we have staff that is dedicated to operate our networks, because even though we do everything right, we have to deal with the over 1,000 pairing partners we have.

Each one of them can leak a route, and bad things happen. What happened?

Okay.

We have over 80 locations worldwide, we're growing that fast. Every one of our sites has somewhere from a few medals to a few hundred medals that we operate. And everything is done over any cast, as I've said before. Our biggest experience that set us apart from others is we deal with DDOS all of the time.

We have lots of customers that I would classify as unpopular. Therefore, we get all kinds of traffic. We have extensive experience in dealing with it. So yeah, less than one percent of the DNS packets that hate our network, or sites, are answered. Not too many people complain about this, because most of the things that we absorb are attacks.

And we have spent an awful lot of time on figuring out how to do a defense in depth, for what we need to do. And we want to

answer the good queries, we want to drop all of the bad ones, we even drop all of the answers, everything.

We also are very adamant that we should not be an amplifier, against others on the Internet. So we have gone out of our way to make our answers as small as possible. And yes, next slide please.

Sorry I didn't know what format to use for the slides. So Adobe would like it, but Adobe will never like anything, so who cares? We have two main DNS services. One is what I call our DNS, which is a standard authoritative server, accept it doesn't have zone files. It has zones, but it doesn't have zone files.

We have DNS infrastructure built in, and we signed every answer on the edge, accept for DNS key records and CTS records. When people enter the [inaudible] API or UI. And we distribute it via our own mechanism to the edges in a few seconds. Our servers are really fast, they're reliable, and one of our latest innovations is that we don't answer any queries with big answers.

If you ask any of our servers for any, you get a little H and four record, and if it's from a signed zone, you asked for signed date, you get a signed answer. We can provide this service for anybody who wants it.

In there is a picture of my other DNS offering, which is what we call virtual DNS. That is a non-standard, let's call it a proxy that sits in front of your DNS servers. Hopefully hidden servers and we ask, when we don't have the answer to any cast, we go and ask the hidden masters what the answer is.

This allows you to operate your own infrastructure. It gets totally hidden by us, we absorb all of the attacks, and the only cost is what you pay us and a little bit of a delay when we have to fetch the answers. But if you answer them out of cast, we get them much faster to you. Okay.

Sorry for that, hopefully we can all see the slide and the presentation when it gets made available. Everybody here has been talking about the cost of staying alive, or online, which is about the same thing these days. I can talk all about how big of an attack we have absorbed. Well, the number in front of the Xs changed last week.

It went from four to five, because we got a bigger attack then we had seen before. At the bottom of the screen, you see an image that I took from our webpage of the current attacks when I made the slide. This is low. This is low.

And there are 19 different attacks that were going on against UDP and there were 20 going on somewhat against the same

target on TCP at the same time. Yeah, so the other thing is, if you want to do your own any cast operation, you have to…

You have to invest in your staff, you have to invest in monitoring and all of that with an outside vendor for full or partial replication of your DNS servers. You can get rid of that headache. It costs a lot of money to operate an any cast network. Dealing with 1,000 plus pairing partners that we have takes a significant amount of time for our team…

The DNS products, we have a virtue on the, so when we have the virtual DNS product, then your servers are protected by our frontend and they're not protected if they are not behind any any cast protection. Go to the next slide please.

I covered that one, next one. Yeah. We have asked, if you want, to talk to me, fine. We have an extensive DDOS protections in-depth. We can do DNSSEC for you, and I want everybody to join us in the fight to get DNS for smaller…

Any cast with lots of node renders it not necessarily more to have 10, 15 NS records. String DNS, they get big. Especially with glue, suppress any, replace RSA. Yes.

And yeah, use any cast.

JACQUES LATOUR: And get rid of IPv4.

OLAFUR GUDMUNDSSON: Yes, I have to stop Patrik. How good is your address data? How do you know that the address you think are in Hong Kong, are actually in Hong Kong, and how many of them are in the U.S. which are actually somewhere else.

PATRIK FÄLSTRÖM: It's actually kind of, it's not as hard as people think. Because if it is the case that people sort of discover that things are on the wrong place, they call us and compared to some other any cast survivors, it has actually happened. People from Chile asked us where we thought Beijing exit was.

OLAFUR GUDMUNDSSON: Okay. Because I've been looking at all of these address maps and they are really…

PATRIK FÄLSTRÖM: No, you're absolutely right. And this year's parties that know you cannot know, and specific now when the depletion of [inaudible] address space. You see these kind of, sort of, route announcement others like all over the place. So first of all, there

are two things there. First of all, this is one of the reasons why it's so important.

It was so important to get the root zone signed, and also others. And also that people start to validate. So basically, as soon as that is done, it doesn't really matter where you get the data from. Secondly, one of the most important things that people…

Sorry, let me take a step back, your responsibility as an any cost provider is not only to respond to DNS queries. It is also to keep track of your routing, and also to answer the phone if it is the case that there is routing [less?], even though people are playing around the IP address.

Of course, many people are to be able to get better services. They are, of course, like having other copies of their own servers or caches, or they put in sort of proxies like what Olafur was talking about, for the IP addresses that we are sort of, the routers that we take care of.

And that sort of works until you see route leaks. At that point in time, people call us, and you as an any cost provider need to be able to answer the phone and keep track of that. So the important thing is just to do the right thing.

JACQUES LATOUR: So I guess we have time for questions now. So Robert.

ROBERT: I have a question about the virtual DNS thing. The cast data you serve, if it is DNSSEC signed, and the [RO Sig?] you have the time as well, how does that work out?

UNKNOWN SPEAKER: The customer defines how long we can cast it.

ROBERT: And then you countdown the TTN as well, and serve so how do you…? So he tells you and on a page that he uses to setup the zone, says that you can catch it for an hour.

UNKNOWN SPEAKER: Yeah. If they're willing to do that, yeah.

ROBERT: Okay.

UNKNOWN SPEAKER: If they want 30 seconds, they can do that too.

JACQUES LATOUR: Any questions? So we've got a couple of minutes left, maybe we can talk about the most designed, most important design

criteria for your any cast platform. Some is number of nodes, some is the size, some is the ability to stop traffic, ability to respond. Like what's the key thing for your architecture?

UNKNOWN SPEAKER: People. It's the people who operate your system, and the systems they put in place. If that's not working well, you know, the…

UNKNOWN SPEAKER: To pick just a single, maybe the diversity to maintain as much diversity as possible, but still think about people that managed this diversity so because managing the diversity can be a problem that operate this to learn all of the different software, different hardware. So that's easy.

JACQUES LATOUR: Andrew.

ANDREW SULLIVAN: Thanks. So actually I think there are, well, either two or three, we'll see, things that I think are really critical here. So first of all, I'm not sure there is one… I'm not sure I would say there is one criteria, because it really depends on what your goal is, right? And in some cases, what you need to do is sink a lot of traffic. I

mean, we saw in the F presentation this morning, that essentially what's happened in the root system is that we're doing, we're now adding any cast nodes essential as sacrificial nodes.

And that's really all it's for. There is no speed addition, right? Because the root is not queried that often. So that's, I mean, that's the kind of thing that you've got there, but other cases what you really want is latency. And in that case, what you need is the number of nodes because you need to be close to a lot of people.

Another really big factor is picking, and I think this was part of the point that Woody was making earlier, you really need to make sure you're picking your network location, and also I guess Patrik, that you're picking your network locations so that you're not defeating yourself, you know, that you don't have bad transit or whatever.

And then finally, I want to pick up on something that Patrik said, about route leaks. One of the things that we discovered, so for those of you who don't know, we bought [inaudible] a couple of years ago, and… So we have their data before, but now I get to see it on the inside. And it's astonishing the crap that's going on out there.

And if you don't have automated detection of problems in your system, just going to be in a world of pain. So a very big chunk of what makes for a successful deployment is good tools. You've got to have good automation for deployment, and you've really got have really good automation for management and understanding of the network, or you're going to be in trouble, because you can't actually run a system by waiting for somebody to call you.

By that point, you're already outside your SLA. You really need to react way before somebody else does. So those are three things. I guess it was three.

UNKNOWN SPEAKER: I think there is a lot of stuff that I would throw into the big category of general competence, and understanding the basics, and so forth. I think once you get beyond knowing that you can do this and just get through a regular day. After that, it's all about scaling and being able to keep ahead of growth requirements is where almost all of the work winds up going, and you know.

Keeping enough people working on the deployment queue and having enough people who are handling customs clearing, and you know, making sure that you've got enough equipment in the delivery pipeline, and with all your provisioning people out there

making the phone calls they need to get all the new cross connects to, you know, so on and so forth.

One of the things that makes a big difference is making sure you're always running dark fiber rather than [inaudible], so that you don't wind up having to upgrade those constantly also, so you can just do an upgrade by agreeing with your peer, person on the other end of your fiber to upgrade your optics and your own port, and you're not also dealing with a carrier in between.

I don't know. Just, I think, peering rather than transit is kind of a necessity in order to grow beyond a certain size. That if you're paying somebody else retail to move the bits for you, you're never going to be able to keep up with what your customers need in the long run.

So you've got to be making your own bandwidth in the same way that, you know, a large Internet service provider, an access provider would for exactly the same reason.

JACQUES LATOUR: Patrik?

PATRIK FÄLSTRÖM: So, peer not transit, really, really important because of the reasons we just heard, over and over again from many of us. You

need to watch your IP addresses and how those route blocks are actually announced, taking track of those. And as Andrew said, not only of course, answer the phone when someone is calling, but you need to watch that.

So as a root server operator, one of the most important things is actually to keep track of your route announcement, where that is visible around the world. And act accordingly. And basically I would say those non-DNS issues are actually the most important ones. The DNS part is easy.

JACQUES LATOUR: Chris?

CHRIS GRIFFTHS: I would agree with everyone here on pretty much everything they said. I certainly echo people first. It takes, you know, I agree. DNS is fairly easy to run, although when you add the complexities of IPv6 and DNSSEC it becomes exponentially harder, and we've seen lots and lots of people make mistakes there. I would say, you know, having people that are designing and building and continuing to scale networks, and continuing to look at that, good measurement and monitoring platforms are absolutely critical to make these deployments successful.

And having people that know how to plan and design these properly is critical. So if I had to say anything, I think having the right people that understand how to scale platforms is critical for this.

JACQUES LATOUR: All right. Any more questions from the audience? So one observation I made from this panel, RIPE Atlas seems to be pretty good tool to monitor DNS infrastructure, the any cast for multiple site to [inaudible] 12. I make a lot of use for dot CA to monitor trace lot graphically.

It shows a picture of where the traffic flows and you can see graphically anomalies in traffic, and then you can work with your provider to fix that. So that's pretty cool tool to play with. I think you have to be a member to use the graphic portion of Atlas to create this.

I'm not sure, yes. So I [inaudible], but that's a great tool to use to see graphically the [inaudible]...

UNKNOWN SPEAKER: I heard those people had spare probes to hand out, so they only need one probe and connect it, and then they can do their own [inaudible].

JACQUES LATOUR: All right, that's it.

PATRIK FÄLSTRÖM: When you talk about probes, think about what Bill and I were talking about, sort of the defense between peering and transit, that you might get problems if you are having transit between the DNS server, and some from the IX. You of course, might have the same issues on the probe side.

EBERHARD LISSE: There are no more questions?

UNKNOWN SPEAKER: Yes, we have a remote question from [inaudible]. "How would you optimize traffic originating from a large CDN network when you don't know the Internet topology of that? Or don't you care? Even if it may affect your service."

How would you optimize traffic originating from a large CDN network you don't know the Internet topology about? Or don't you care? Even if may affect your service.

UNKNOWN SPEAKER: I'll try. Let me try simplifying the question first, and seeing if that, maybe they can say whether they [inaudible] of it. So if you have a CDN, like a DNS network is a content distribution network, you have content that you're serving out.

How you originate traffic out of it. How you specify the routing out, is largely dependent upon where the queries come in. And that can be true for any kind of CDN. So if you're hosting web traffic, you answer typically from the place where the query came in. The exception to that would be load balancing or internal failover. So if one of your servers went down, and you have a way of redirecting a query in flight to another server, or if one of your servers is overloaded and you're able to load shed to another server, then that server might originate the traffic out instead.

Typically there is also a layer separation between the servers, which are creating the answers and slapping on an IP address on them, versus the routing which is receiving a packet with an IP address on it and forwarding that on. So, the routers don't really care whether it's a response to a query or a query itself, or something else entirely. They're just forwarding packets based upon the IP addresses.

So the routing policy is done in the routers, but the application layer can put a different IP address on something. In a stateless

protocol like DNS, the only IP address you would ever put on a response is the originating IP address of the query. So you don't actually have a handle on anything there.

It's not, that's not something you would fine tune or change.

PATRIK FÄLSTRÖM: So the, I think maybe the question might also be about, for example, if you have a large ISP, which have a natural recall over the place, and then you have a resolver which is the one that actually sends the query to the authoritative server, so that might also be the question which means that it would be tied to for example the various work in the IETF regarding where the client is, and [inaudible] zero sort of extension stuff.

What we have to remember here is that the response is supposed to be sent back to, as Bill said, to wherever the query is coming from, which means the distribution is actually more a question of the CDN. And then the one that is responding to the authoritative server.

That said, if it is the case that there are parties that do send a lot of queries, for example, to us. Then of course, we attack that and we have some communication with them to see how we can optimize and ensure that the traffic actually works as smoothly as possible, even under a denial of service attack, etc.

UNKNOWN SPEAKER: That's it. That's all. All right, thank you all. [APPLAUSE]

EBERHARD LISEE: Okay, thank you. And now Jacques will give us some closing remarks.

JACQUES LATOUR: All right. First time, I'll give it a try. So where is Andre, if I'm not doing good, go like this. I'm doing okay, go like that.

So I guess I'll recap the day, right? So this morning we had a presentation from a good presentation from Andre about not DNS [inaudible]. So that was interesting. I learned a lot about email address internationalization from Woo I guess.

Marvin. The issue is we have IDN for dot CA for French, but none of the browsers. We have issues in using IDN in Canada with French character, and all of the applications, so that's something we're looking into. We had a good overview of the overall process. Francisco did a good job there, I guess to understand, provided a better picture on what, how it works and in an actual implementation of that.

From Turkey, Atilia did a good presentation on sharing details of the attack. I thought the lessons were pretty good in there.

There is a lot of issues we need to look at and address, and certainly having IANA respond urgently when you're under stress, would be a good future to have.

Maybe a special, special hotline when something bad happens, that would be useful. TLD ops did a, Christian did a presentation of TLD ops. It's our ccTLDs to use, to share to security contacts. There is a lot of information that can be shared there. We're not using it enough today, but we should be thinking about that, if somebody gets DDOS or something, they should email to TLD ops to say, hey I need help.

I need advice. And I think there would be a lot of that information in that group of specialized security people. So it's our to use. Japp, as usual, he makes sure that the Internet is up and running and he's got all of his monitoring tools. So that's not the first time with his presentation on DNSSEC.

F-root overview from Vicky, that was all right. Then we had universal acceptance. That goes with the IDN email. There is a lot of work to be done. I think it's a mindset that has to be put in place for young software developer to think about that today, and ongoing.

So there was a discussion that there was a future for new products, so it needs to be core to a product. So universal acceptance is something that we really need to look at and

make sure that we pay attention to. I thought Richard was doing something interesting with monitoring of his any cast, with…

So I would like to see further on where that goes on the right Atlas platform. And that links in to the panel where there is a lot of monitoring for any cast that we need to develop or make easier to use, or understand.

I think it's very difficult to get the real time snapshot of your DNS any cast platform at any moment in time. So if we had really cool tools that could give us real view so that people can manage any cast better. You get better visibility. You can understand where traffic flow then it would be a better world. That's it.

EBERHARD LISEE: Thank you. Okay, thank you very much. I'm quite happy that so many people stayed on until the end. Usually more than the ccNSO meeting, which is a good thing to have. All right. I'll see you in probably Helsinki.

[END OF TRANSCRIPTION]