ICANN71 | Virtual Policy Forum – Tech Day (1 of 2)
Monday, June 14, 2021 – 14:30 to 16:00 CEST

KIMBERLY CARLSON: Hi, all, and welcome to Tech Day at ICANN71, Part One. My name is Kim Carlson and along with Kathy Schnitt we are the remote participation managers for this session.

Please note that this session is being recorded and follows the ICANN expected standards of behavior. During this session, questions or comments will be read aloud if submitted within the Q&A pod.

If you would like to ask your question or make your comment verbally, please raise your hand. When called upon, you will be given permission to unmute your microphone. Kindly unmute your microphone at this time to speak.

All participants in the session may also use chat pod. Please use the dropdown menu in the chat pod and select respond to all panelists and attendees. This will allow everyone to view your comment. Also note private chats are only possible among panelists in this Zoom webinar format. Any messages sent by a panelist or a standard attendee to another standard attendee will also be seen by the session hosts, co-hosts, and panelists.

Finally, this session includes automated real-time transcription. This transcript is not official or authoritative. To view the transcription, click on the closed caption button in the Zoom toolbar.

Thank you all for joining, and I'll turn the call over to Dr. Eberhard Lisse.

EBERHARD LISSE: Hello, and welcome to our—I haven't counted how many Tech Days it is. Again it is virtual. I am sitting in my practice in Windhoek where it's afternoon and quite cold. We were supposed to be in The Hague where it's supposed to be quite warm, but we'll see whether this is going to happen next year.

As usual, I'll go a for little bit through our agenda. We have as it is a policy meeting a very short agenda because we share the day so to say with the DNSSEC group who does their thing in the morning. After I'm done with my opening remarks, Jordi Iparraguirre from .EU will talk a little bit about their abuse detection and mitigation.

Then Cristian Hesselman will speak about some project they're having at the SIDNL about the Distributed Denial of Service (DDoS) Clearinghouse. And he will also give us a few slides about his organization. We usually do a host presentation, but we don't really have a proper host this time being virtual. And [ISDN] is supposed to be the host. He was graciously available to put a few slides to show us what they are doing and how big they are.

Then Ray Bellis from ISC is going to show us a little bit about his graphical Atlas monitor. Most of you know what an Atlas probe is. Jaap Akkerhuis has on several Tech Days handed out a few probes. Ray has written a little tool where you can see them on a browser.

Then between blocks we'll have to have a little break of about 30 minutes. If you can give the next slide, there you go. After that, Jothan

Frakes is going to give us a little bit of an update of the public suffix lists. We have heard him before; we will hear him again. It's always good to see what's going on. I've heard that he was a little bit under the weather today, so it may well be that we have to move Howard Eland's presentation about how Donuts is changing everything away from SHA-1 forward. And then Jothan may do it afterwards, or if he's not feeling well, then we will just stop after that.

If you can have the second slide again that I can see, it's Jordi who is going to be next. We basically…I don't know whether we want to see the face of the presenters doing the presentation. I'm going to turn my screen off because it's taking some bandwidth. And we can then turn the screen, the TVs on at the discussion.

One thing I forgot, I have volunteered Jacques Latour to give us the usual closing at the end.

Without further ado, Jordi, you have the floor.

JORDI IPARRAGUIRRE: Thank you very much. Let me share my screen. Okay, the host has disabled. May I have the screenshare, please?

KIMBERLY CARLSON: Jordi, one moment, please.

JORDI IPARRAGUIRRE: Yes, thank you.

KIMBERLY CARLSON: Jordi, you're a co-host now.

JORDI IPARRAGUIRRE: It's enabled. Thank you very much. So there we go. Can you see it? Can you see the screen properly?

KIMBERLY CARLSON: Yeah, we are seeing your presenter view and not the full screen.

JORDI IPARRAGUIRRE: Okay, wonderful. So it is the wrong screen. Right.

EBERHARD LISSE: You have unshared the screen from what we can see here.

JORDI IPARRAGUIRRE: Yeah, because I'm disconnecting the other monitor, and let's see if it works now. Okay.

KIMBERLY CARLSON: Jordi, I have it queued up as well if you would like me to share.

JORDI IPARRAGUIRRE: Yeah, just one second. It does not work like that. Okay, where is the slides? Okay, here it is. Share. Okay, does it work now?

EBERHARD LISSE: Yeah, but we see you're sharing the wrong screen. You're sharing a terminal screen.

JORDI IPARRAGUIRRE: Okay, wonderful. So, yes, Kimberly. If you please can share the slides.

EBERHARD LISSE: There you go.

JORDI IPARRAGUIRRE: Okay, thank you very much and apologies for the technological glitch. So very briefly I'm going to talk about and present very briefly what we do at EURid to detect and mitigate fraud online. Basically, [we are on that] because we care about our users, the businesses, the consumers, the people on the Internet in general. And at .EU, at EURid, we try to do the best possible to make that possible, the have a trusted .EU space for everyone.

But then we have to move inside a kind of framework so basically, as you know and many of the registries do, we do not host content. We do not want nevertheless to be the resource as a registry to be used causing harm to other people.

In addition, we are not the police, not the judge of the Internet. So we may think something about the content, but that's not our business. The only thing we do is, in fact, to verity registrants' WHOIS data. That's

what [we're entitled] to do. That's the mandate we have, and that's what we stick to. But nevertheless, can we do something extra? Can we go a little bit further without getting out from that framework to really help people, to prevent and mitigate abuse?

Well, we think we can. The point is that if we are not the police, not the judge of the Internet, how can we prevent that? I mean, at the end of the day, what's abuse? There are lots of discussions even in ICANN about what's abuse on the Internet, DNS abuse and all that. So let's try to focus a little bit on that.

This is a real example of a clone of the bank that we found on a domain name that had nothing to do with the bank. Not because of their registrant data, not because of the domain name itself. And then we double checked that with the bank and, of course, it had nothing to do. So there are certain things that are crystal clear that that's not fair. Where we would never recommend one of our friends, relatives, whatever to use one of those sites. So for us, this is an abuse, right?

That's another one. This is a clone of the French government page to deal with the taxes. You can be reimbursed or you can pay the taxes. And again it was in a domain name that had nothing to do with the French government, so it was an exact copy of that page.

Another one about pharmaceutical products, some of them that usually need a prescription. But then the associations that deal with pharmaceutical products, we're in touch with them on that but they're [inaudible]. These things are usually fake or stolen products, and you

never know about the composition. So that's not fair. That should be reported somehow.

Then again about very well-known brands with huge discounts, and when you talk to those brands this webpage has nothing to do with me.

So while here we have just some examples of webpages that are clearly an abuse and that can harm people, can harm the users of our domain. So the point is that, yeah, we more or less have an idea of things that we may be looking for. So let's first talk about the concept and then we will move about how we do it.

EBERHARD LISSE: Jordi?

JORDI IPARRAGUIRRE: Yes.

EBERHARD LISSE: Can you let Kim know when to advance slides? We were still on the first one.

JORDI IPARRAGUIRRE: Okay, yes. Sorry. So, Kim, please, can you move forward two slides?

EBERHARD LISSE: Right now we're on Slide 1, Prevention: APEWS.

JORDI IPARRAGUIRRE: I cannot see my Zoom now. Well, can you please go to the ING page, the Number 5 on the stack, please? Thank you.

KIMBERLY CARLSON: Yeah, one moment.

JORDI IPARRAGUIRRE: Yes, thank you. Okay, so this is one of the examples I was talking about. Next slide, please. This is another one, the clone of the French taxes webpage. Another slide, please. The pharmaceutical one. And then the next one about very well-known brands. So now next slide, please. Prevent and mitigate abuse, the concept. Next slide, please.

Basically, the idea is that we act at different levels. At four different levels. Next slide, please. The first one is prevention. As soon as a domain name arrives, we try to prevent it to get into the root if we think that's suspicious. And this is the well-known APEWS. You know APEWS, but I'm going to refer to that later on. This is a system that decides based on different data if a domain has chances to be suspicious. Next slide, please.

The second one is detection at post delegation. Once a domain name is already on the root, once a domain name can be used, we try to detect domains that can be abusive. Next slide, please.

Then in any case we do checks. We have different ways of checking. Basically, as I mentioned initially, we care about the WHOIS data so basically we have the process that's called the WHOISQuality process in

which we ask the registrant to validate the WHOIS data. And then we just introduced the Know Your Customer (KYC) one that I'm going to refer to later on. So we have checks to really see if that owner, the registrant, the person, or the company behind the domain name is really that person. So to know the customer. Next slide, please.

Finally, we have the communication. Whatever we find [around] and we think that can be suspicious as we are not the experts, we share that. Next slide, please.

Here you have the complete picture of the concept, the idea. Some things are done before the domain gets into the zone. Some of the things are done after the domain gets into the zone. And in any case, we share with professional partners about cybersecurity. Next slide, please.

So the practice. That's the concept; now let's get into the practice. How does it work? Next slide, please.

The first one, the one I mentioned, the pre-detection is APEWS. That's an acronym for Abuse Prediction and Early Warning System. That's an award-winning machine learning system that was developed together with the University of Leuven in Belgium. Basically, it works before the domain is delegated. So the domain is registered. It belongs to that person, that company, but it's not yet delegated.

Using different machine learning checks based on data that's provided by security companies, so professionals of all that, the system just

comes out with a [flat out] yes/no. This may be suspicious for future activities or no. Next slide, please.

Basically, the point here is that the domain name delegation is stopped if the system thinks that it can be abused in the future. Nevertheless, we always have humans, we always have people reviewing the case. So the artificial intelligence system may decide something, but that's not going to be never, ever the final decision. There are always people behind reviewing the case.

And now with the Know Your Customer system I am going to refer to later on, we automate this and increase the certitude of that. The Know Your Customer process works.

APEWS took some time to develop, but before that we had to negotiate with the European Commission to change the regulation of the .EU to be able to do these kinds of checks. Because initially our regulation, our legal framework did not allow for that. So it's in production since January 2020 and was in test in 2019 and so on. Next slide, please.

APEWS basically uses different machine learning algorithms to try to identify these kinds of patterns. It retrains itself regularly because, of course, the environment changes so well we could be missing important points there. There is a set of different academic papers that if you are interested in knowing the details about how it works and performance and all that, you can get them in that URL.

And beside and in addition of APEWS we have another system that's based on rules. That was used, for instance, when the European

Commission asked us to stop the delegation of domains that were related with COVID-19. The Commission was conscious that some of those domains could be used to trick people, to sell masks that were not safe or medicines or whatever that were not really correct. So they wanted us to prevent the delegation, not the registration. So that's what we did. Next slide, please.

So that's a picture where APEWS fits. We have the dark blue square, and APEWS is at the very beginning and then raises a flag. If everything seems to be okay, the domain gets into the zone. If it's not, it gets into the legal department. Next slide, please.

The second block is the post-delegation checks. The domain is already in the root. It can be used. So what are we looking for? Well, basically, we look for the domains that got delegated in the last 24 hours, the domains that were delegated some weeks ago because you have to allow certain time for registrants to have content, to check the WHOIS data or whatever. And also, we can feed the system with a specific list of domains at any time. All the domains that start with A, all the domains that have dot string, or whatever you want.

The idea there is—next slide, please—that we analyze the domain name itself. We may look for certain brands, certain keywords, certain specific things that based on our experience we think can be interesting to monitor.

We crawl also all those domains and then we analyze the web content, the HTML or the JavaScript and all that. So is that a web shop, for instance? If it's a web shop, what is that selling? Is that selling well-

known brands with huge discounts? Is that selling medicines? Is that a copy of a bank or whatever? So we can create new modules and add them to the system to identify these kinds of things that we think—and it's just our opinion—that could be strange. What I said initially. We are not going to recommend that to anyone. So that's what the system does, and we analyze that together with certain data and metadata of the registration. Next slide, please.

That again, the dark blue square on the right, so the domains are in the zone and work on those domains. And whatever the system may think, again, is sent to the legal department. Next slide, please.

And that's where the checks happen. The third phase of the concept map initially is that where the human really intervenes. The artificial intelligence system, the rule-based system, whatever, may just send you a list of different suspicious sites, but then we have people who really are reviewing that.

And in any case initially what we do is that we start our WHOISQuality checks. That is that we ask the registrant to prove that the WHOIS data is real. So that creates a lot of workload on our side. We may also be receiving forged copies of ID documents or whatever but, again, we are not the experts on that. That's not our business. That's not our core business, so that's why we just implemented the Know Your Customer technologies that you are going to see in the next slides in which it's much more difficult to really lie, to really trick the system there.

And again, whatever the artificial intelligence system says, whatever the human thinks about that, we are not going to suspend the domain,

the registration based on our opinion on the content. We can just suspend the domain names if the WHOIS data cannot be proved by the registrant. So what we do is that we share our suspicions with experts. Next slide, please.

Here you have an estimation or an idea of the workload that that represents to the legal department. In red basically you have the suspended domains or domains that were reviewed and because they did not answer to the WHOISQuality requirement or Know Your Customer requirement had to be suspended. So there is a lot of workload there. Next slide, please.

Here is where we have them, the legal department in the middle of the whole process, the humans, the people there interacting with the registrants, with the registrars, with the courts or whatever if need be or the other way around. The prosecutor and the police can go to us asking to review certain documents [and so on].

So if that registrant cannot prove its relation with the WHOIS data provided, then the domain is suspended. Otherwise, whatever the kind of content, we are not going to suspend the domain unless we receive an order from a justice or police [or so on]. Next slide, please.

As I mentioned the fourth block is we share, we communicate, we share intelligence. So basically as we are not the experts on that, what we do is at the same time that we are going to start the WHOISQuality process to check the identity of the person that's behind a registration, we share that domain name with different entities.

Here just to name some of them: Europol, the Belgian CERT and Center for Cybersecurity, the Belgian customs, the professional Alliance of Safe Online Pharmacy, and so on others.

So the idea is that we do not know. We think that something is wrong there, so let the experts do their work on that. So it's a kind, again, of an early warning. Next slide, please.

That's important because the point is that we want to do that and we do that daily and in parallel to our own processes. Because—next slide, please—the sooner we inform to those third parties, to those professionals the faster they are going to be able to cross that data with other sources to classify that kind of issue and then to put up a [barrier] somewhere, for instance, on the safe browsing list. So whenever a final user is trying to access that domain name some professionals of cybersecurity have reviewed that previously and then they will, for instance, make it difficult for the Internet user to really access that domain name.

So on our side, basically, we just raise a flag and say, "Hey, wait a second. We think that this is dangerous, but let the professionals decide—the police, the Europol, the cybersecurity centers, and so on," and then they will do their business. We do our part. We look. We check. We check for the WHOIS data quality. We inform. And then the rest proceeds. Next slide, please.

And then, yes, from time to time you do not just get success but your success is acknowledged. Here is just an excerpt of a news item that's signed by the financial and economy ministry of the Belgian

government. So, yes, sometimes these things go public and we may help to prevent abuses online. Next slide, please.

Nevertheless, this is a Darwinian marathon. This is never going to end. We have started to push in one sense. We are looking for the quality of the WHOIS data. And then you start realizing that you are starting to get nice WHOIS data but completely nonsense. Because it's very easy to find John Smith living on Market Street Number 7 and equivalent in other countries and other languages and all that. So apparently may look good but at the end of the day this is of no use whatsoever.

So we push, they adapt to bypass our detections or to just go away to somewhere else. But if they change, then we have to adapt or we fail. So this is a kind of…it's very balanced. Sometimes we may think that look at that, we came, we are able to expose some kinds of abuses from our TLD. But for certain cases, you will never be sure. You will never be sure because is that them that left or is that us that we are not able to catch them and find them? Next slide, please.

So we've seen changes of those behaviors. We've seen some abuses that have disappeared from our registry. Some of them it's clear they disappeared because, for instance, certain patterns and certain users of domain names we don't see them anymore. And the domain names are easy for us to check so, yes, these kinds of abusers just disappear.

But other abusers, they just also disappear but are they hiding somewhere else in our registry? We don't know yet. So that's why we need to reinvent ourselves and look for new ways to be sure that we can offer a space as safe as possible.

The point is that, as I mentioned, it's very easy to fake the WHOIS data. John Smiths everywhere and all that. In addition, we also have anonymizer services. Some are companies. Some others are coming from the registrar itself. The point is that as we are playing only on the European Union area, the GDPR is to the rescue. So there is no point in really using those services. But nevertheless, they are there being used and we are starting to see that some of those abusers are hiding there.

So then it's the final thing that we put in practice last April, it's the Know Your Customer initiative. With the Know Your Customer initiative we are starting to use eIDAS which is the European electronic identification system and services. So for the citizens that are having the electronic IDs from the European countries, we can ask them to identify themselves using that. So that identification is going to be clear. There is no way to forget that.

And another one is to use what's called an MRZ area on your passport or your ID card. It's that part of the passport or the ID card in which you have characters that are all using a kind of, let's call it, computer form. That can be readable. It's MRZ because it's machine readable zone, so the machines can read that. And then by ourselves getting a picture of that we can double check the content of it and be sure or not if that person is faking or is not faking an identity using that system.

So that's what we just started to offer to implement in April, and we are working on that to counteract some of those abusers that are just using anonymizer services or names that can be valid but, in fact, are not really the ones of the registrant.

So the point is that we are not using that to everyone. What we are looking for is abuses and then we start those checks with those that have suspicious registrations. Next slide, please.

So that's basically all on my side. I'm just going a little bit fast to allow some time for questions if there are any. Thank you very much for your attention and time.

EBERHARD LISSE: Thank you very much. Interesting presentation. Jacques, I see your hand. You'll get in just now. Interesting presentation. I like the idea of, especially since you fall under the GDPR, to allow everybody but forcing them to identify themselves.

There is a German ID card that works. I know the Belgian ID card is also working on a large number of electronic systems. The Dutch one I don't think is working. And the Estonian one, especially the e-Residency thing is also working. So that's maybe the people who do not have one can relatively easily get a verified ID card. You have to identify yourself to the authorities if you get such a card.

I think it's a good idea to know your client even if there is no legislation forcing you. There is no legislation prohibiting from this, and there is also no legislation prohibiting from saying this website looks suspicious. We report it to somebody who can look at it, as long as you don't act on the registrant unless they're at fault.

Jacques, you have the floor.

JACQUES LATOUR: Thank you. Jordi, quick one. You mentioned a couple of times the acronym KYC.

JORDI IPARRAGUIRRE: Know Your Customer.

JACQUES LATOUR: Know Your Customer. Okay, maybe you should add that as a little bullet somewhere to explain it.

JORDI IPARRAGUIRRE: Yeah.

JACQUES LATOUR: I like that. Know Your Customer. Thank you.

JORDI IPARRAGUIRRE: We use it so often that it is becoming part of our language. Sorry.

JACQUES LATOUR: Yeah, I get that.

EBERHARD LISSE: KYC means, obviously, know your client, and that's a bank standard that is in financial transactions. It's due diligence, that kind of thing, looking in financial things with the ultimate beneficial owner. But the

banks require nowadays for every account holder to be properly identified. And once in a while we see this here. Now maybe they lost the papers and we have to resend them, but they want to know whether the person you're dealing with is the person you're actually supposed to be dealing with.

In Namibia it's mandated by law. In many countries it isn't. For companies for registries it's not mandated. We also like to know our client, so we ask also for some form of this. As long as you don't publish the information in violation of the GDPR, I think it's a great thing.

Cristian, you are next. Jacques, you can take your hand down, please.

CRISTIAN HESSELMAN: Yeah, thank you, Eberhard. And thank you, Jordi, for the interesting talk. I was wondering, do you guys also share with your registrant why their registrations have been flagged by your algorithms?

JORDI IPARRAGUIRRE: No. Basically, because there is a risk there. If they are going to…in fact, if this is an abuse, we are going to give them hints about how we detect those abuses. What we do is though we share that internally so the legal department knows that has been flagged because of, I don't know rule one or reason whatever or all that, which is what the system thinks may be behind that. But at the end of the day it's not really critical because they just process the data and we focus not on the content but on the WHOIS data itself. So the address looks incorrect or, I don't know, the

address does not match the format or the language of a given country or whatever. But not with the registrant.

CRISTIAN HESSELMAN: Okay.

EBERHARD LISSE: I see….

CRISTIAN HESSELMAN: That's interesting because we…sorry.

EBERHARD LISSE: Go ahead.

CRISTIAN HESSELMAN: Yeah, that's interesting because our philosophy at .NL is slightly different. We do share that information. And the reason is that we just want to increase the bar for the bad guys to make use of .NL domain names. So if we also…we don't have to be very specific about what you do, but you can publish what kind of indicators you look at. And then let's say the threshold increases for the bad guys to do something with .NL domain names. So that's our philosophy which is slightly different, I guess.

JORDI IPARRAGUIRRE: Mm-hmm. Yep, that's interesting.

CRISTIAN HESSELMAN: Interesting to hear. Thank you.

JORDI IPARRAGUIRRE: Yeah, thank you very much.

EBERHARD LISSE: I mean, publishing your root is one thing, but informing registrants that they have been triggered and then in the end they pass it anyway, I agree that would not make sense. I think the answer is somewhere in the middle. What Cristian says is important, but I also agree if your thing gets flagged, if then the human review decides it's valid, there is no need to inform the registrant.

Let me go the Q&A pod. There is one question from John McCormack. Did it cope well with the Brexit re-registration of British and Northern Irish .EU domain names or were they flagged as abusive?

JORDI IPARRAGUIRRE: Let me read it again.

EBERHARD LISSE: Did it cope well with the Brexit re-registration or British and Northern Irish .EU domain names or were they flagged as abusive registrations?

JORDI IPARRAGUIRRE: All the domain names that were related to Brexit were managed differently in another flow, in another process and completely following another kind of rules. So we just focused on them and we treated them differently.

EBERHARD LISSE: And then there is a question from Syed Shah. Can you explain a little bit about rule based system that you are using? For example, currently as you said we can stop COVID related domain names.

JORDI IPARRAGUIRRE: Okay, we got the requirement from the European Commission last year, last spring when COVID was just ticking up and all that and it was clear that we had some places that were selling tests that were noncompliant or masks that were not verified by the medical partners or the ministries or whatever. So there was a risk of harming consumers, so the Commission gave us a list of different words that we have to look for in the domain name. So COVID itself or mask or virus or that, whatever, in all the different official languages of the European Union.

So what we did was to add all those strings in different rules so whenever a domain name was registered and contained one of those strings it was stopped. It was not delegated, and it went to the legal department and the legal department started the WHOISQuality process.

That created a lot of workload because, of course, there is a new trend and then you have the domainers that just want a domain to resell for

a better price later on. You have the abusers. You have the people that are really interested in, for instance, sharing information about that. So there was a lot of demand and we had a lot of work to manage those domains.

In addition, for certain languages you have these strings that we had to stop are part of words. If I remember, I think that one was mask in a certain language was also matching machine in German for instance and virus in one of the languages was just "vir." And of course, there are a lot of words that have the "vir" string inside them. So it was very difficult really to deal with that and manage it properly. But, well, we did it and that's why we have now that system in parallel because we were required to stop the delegation of those domains before getting into the root.

Did you have more questions?

EBERHARD LISSE: Okay, thank you very much.

JORDI IPARRAGUIRRE: Thank you very much.

EBERHARD LISSE: If there are no questions in the Q&A pod, I propose we move to the next topic. That will be Cristian Hesselman from .NL speaking a little bit about the effort to create a clearinghouse and also a little bit about the host. You have the floor, Cristian.

CRISTIAN HESSELMAN: Thanks, Eberhard. Thank you. Kim, can I share my…can I present my own slides, please?

KIMBERLY CARLSON: You should be a co-host now.

CRISTIAN HESSELMAN: Oh, okay. So share screen. Oh, there we go. I pressed the wrong button. Thanks. So this is the one, I think. Yes, here it is. Okay, can you guys see my screen now?

EBERHARD LISSE: Yes, we can. You must just go and [play here] on the slideshow.

CRISTIAN HESSELMAN: Okay. Yeah, right. Okay, good to go. All right, so this talk is about a DDoS clearinghouse. And I'll explain what that is in a minute, but the essence is it's about sharing information about DDoS attacks across different organizations. That's basically the punchline. And this is work that's taking place in a research project called Concordia which is a weird acronym of Cybersecurity Competence for Research and Innovation. It's a collaborative effort.

It's something that we do at SIDN Labs but in collaboration with the University of Twente, Telecom Italia, FORTH (that's a research institute from Greece), University of Zurich, SURF (from the Netherlands),

University of Lancaster, and CODE Research Institute (from Munich, Germany).

I'm getting background noise here because of construction work. Am I still okay on the sound levels?

EBERHARD LISSE: Yes, it's fine.

CRISTIAN HESSELMAN: Okay, good. Thanks. All right, so DDoS attacks, I don't think I need to explain to everyone what that is. But essentially, it's sending a lot of traffic to a single target from multiple sources, or at least from a large number of sources. And we've seen quite a few of them over time. Maybe the one that we know best in the past were the DDoS attacks on Estonia in 2007. Took out many websites, including critical infrastructure such as banks and governments. And then, of course, in 2016 we saw the Mirai botnet. That was really an IoT powered DDoS attack. And then recently we saw DDoS attacks in Belgium, for example. That was in May of this year and also previously in September 2020 on a few ISPs in the Netherlands.

What started this whole endeavor was basically the attack that took place in the Netherlands in January 2018 which took out several banks, several government agencies, and those kinds of websites. And so we thought maybe there's something else that we…maybe there's something more that we need to do as a community in the Netherlands than just mitigating the attack but perhaps we should also start sharing

the information. So that's basically when we set up this whole DDoS clearinghouse initiative, and we also made it part of this European project which is Concordia.

The goal of the work that we carry out in Concordia is, what it says here, to collaboratively protect critical infrastructure in Europe and in the Netherlands against DDoS attacks. The way we do that is through a clearinghouse which shares information about ongoing DDoS attacks across different organizations.

What we want to do in this project is we want to pilot that system. We want to develop it, pilot it, use it in Europe and also in the Netherlands. And also make everything open source and open source design so that it can be used anywhere so that any group of organizations who wants to set up their own DDoS clearinghouse or similar organization can do that.

The key outputs of the work, there are actually three. There are two pilots that we're scheduling, one in the Netherlands and one in Italy, and both will be based on the DDoS clearinghouse software that we're developing here. I'll be talking about that a little bit more later on.

And then we have something that we call a DDoS clearinghouse blueprint or a "cookbook." That's basically capturing all the lessons learned that we have gained over the past two years, and when we're done in the preceding four years, so that folks can really learn or benefit from what we've learned in our endeavor in this project.

So the key innovations are that we want to bridge the gap basically from research to deployment. Which is more than just acknowledging because it's also about setting up the required legal contracts. It's about setting up an organization that will manage this whole system and so forth. I'll be talking about that a little bit more too.

Another innovation is the open source design that I spoke about. So that's the cookbook with all the lessons learned in there using the DDoS clearinghouse in the Netherlands as our main use case. And what I said, we want to enable other groups of organizations to set up their own clearinghouse.

From a technical perspective, one of the innovations is that our system can operate across different types of networks which is important because some folk have networks that capture DDoS traffic through PCAPs and others use Netflow. So that can change and the DDoS clearinghouse needs that kind of information to gather meta data about the characteristics of the attack.

So this is the high-level concept of how it works. Basically, what you're seeing on the right are three service providers, SP1 through SP3. And in this specific example SP2 is getting hit by a DDoS attack. It needs to absorb it, so it needs to have its mitigation services in place be it their own or somebody else's. So when they hire an anti-DDoS service such as Akamai, for example, or Cloudflare or whatever, so that's what needs to happen.

But in the case of the DDoS clearinghouse, the service provider under attack, so SP2, also shares a fingerprint of the DDoS attack with the

other service providers. So SP1 and SP3 in this case. And together these three service providers form what we call an anti-DDoS coalition. So they have a mission to collaborate and to collaboratively fight DDoS attacks.

The advantage of sharing the information with the other service providers is that they are prepared in case the attack comes their way next. So SP1 already knows how to configure their infrastructure based on the fingerprint before they actually get hit by the DDoS attack.

So for SP2 the whole system is still reactive because they need to react to the incoming DDoS attack. But for SP1 and SP3 the DDoS clearinghouse gives them a proactive edge. So for them it becomes a proactive thing. So it basically buys these providers time to properly conference their systems so that they're prepared for the attack.

What's important to highlight is that this is not a system that replaces DDoS mitigation services. So you still need services that scrub traffic or that blackhole it or whatever. This is an additional layer of information sharing on top of that infrastructure. So it's an add-on rather than a replacement, so that's important to highlight.

And the concept is generic. So the anti-DDoS coalition can be service providers within a specific country such as in the Netherlands, for example, but it could also be across member states in the European Union. It could be the business units of a global company. It could be anything. As long as you have a set of organizations or semiautonomous organizations that are willing to share meta data about DDoS attacks to collaboratively fight these attacks.

So that's the key concept. This is an example of what these fingerprints look like. They're being generated by a component in the DDoS clearinghouse called the dissector. I'll be talking about that a little bit more later on. The dissector sits in the network of the target, and the target basically captures information about the…it measures information about the incoming DDoS attack.

So it could be about the type of protocol, the type of DDoS attack it could potentially be like an amplification attack, for example. Numbers of packets, duration, that sort of stuff. So it's basically a block of meta data that's being shared across these different organizations through the clearinghouse.

The clearinghouse is also important to increase what we call digital autonomy because it gives potential victims…it widens the view of these service providers because previously they only saw the DDoS attacks that hit them, but now they also get information from the other organizations in the anti-DDoS coalition. So they get a better view on the anti-DDoS landscape.

And I recently spoke to Dutch National Bank about this specific topic, and that's what really triggered them. So that's something they found really interesting.

Also, it gives organizations in an anti-DDoS coalition more control over how they can handle DDoS attacks. Because they have more insight into the data of DDoS attacks, so they can better organize their systems and processes and it can even be a guiding factor for what services they

want to contract with. So what services they want to buy to actually handle DDoS attacks, for example.

And then finally the concept is also I think good for building up a pool of expertise. So an anti-DDoS coalition builds up a pool of expertise which is independent of specific DDoS mitigation providers. So you basically get more knowledge and more data into the DDoS ecosystem, so to speak, rather than all that information and all that knowledge being with a few large companies that can handle these DDoS attacks.

So this is what the architecture looks like. What you're seeing on the left is an incoming DDoS attack at a victim. That information is being sent, so that the network traffic is being sent to what we call a dissector. So this could be a PCAP capture of the incoming traffic or a Netflow specification. That goes into the dissector.

The dissector creates a fingerprint out of it and puts it in what we call the DDoSDB. So that's a central repository that contains all these fingerprints. And it gathers information from multiple actual victims, and then it sends it to potential DDoS victims. So other organizations in the anti-DDoS coalition. They can request that information from the database and then use it to create mitigation rules that they can put into their mitigation devices. So that's being done under the control of operations [teams.]

And then in the middle there are a couple components that we call the supplementary services. So the dissector database and the converter are what we call the core components, and then we have the supplementary components which basically sit on top of that and make

use of the information that's in the DDoSDB. To run the concept you at the very least need the core components.

So that's pretty much the overview. And then we have one component that's responsible for…I'm not going to discuss this in detail, but this is a component called the converter and it's responsible for creating those rules that you put in your mitigation infrastructure and mitigation devices. And currently it can handle Snort converters and Snort rules and it can handle IP tables, but we foresee that additional extensions could be added here.

This is where we are right now. We've developed all these components. Some of them are at a relatively high maturity level. Some of them are more, let's say, need more work. But at least the dissector and the DDoSDB and the converter, they're in the either high maturity level or medium. And that's important because we want to carry out a pilot with the whole system.

This is a screencast. I hope that works. I'm playing it now. I'm not sure. Do you see something moving? I think you do. So what you're seeing on this screencast is on the right are DDoS tools. I think it's DDOSIM at this point. So they're often used to generate DDoS attacks. And this is a test setup. It's being used to generate the fingerprints.

So in the middle we're going to see the creation of the fingerprints, and on the left we're seeing the DDoSDB. So that's the database in the middle. So when you play that screencast—I'll forward it a little bit because it takes too much time—so they're initiating the DDoS attack. Then at some point—where is it—so now the attack has been

discovered and it's starting to capture the traffic. So they're creating PCAPs.

And than at some point it's being…so now the DDoS dissector is being started and the PCAPs or the traffic captures are being converted into a fingerprint. And that should be happening any time now, looking at the middle screen now. There we go. So that's the fingerprint. It just came out. The fingerprint has been created, and now the fingerprint will be sent to the DDoSDB on the left.

So fast forward again just to show you guys that it's working. Fast forward. So that's at the bottom of the bottom of the screen in the middle, there's the fingerprint and the URL of the fingerprint has been uploaded. And then refresh on the DDoSDB and we see that there's a new fingerprint in the database. So this is basically the main flow of the system.

So then we have another screencast which is this one. This is what we call the DDoS grid. This is being developed by the University of Zurich. That's one of the supplementary components that makes use of the fingerprints in the database. Actually, what's most interesting is the last part of the video. This is where it creates the fingerprints. There it is on the left. And then on the right we're going to see what it looks like.

So it imports the fingerprint into this DDoS grid visualization engine. Then it says visualize, some meta data [that we add] first. Forwarding a little bit, and some more. Then it's being analyzed and uploaded into the DDoS grid, and finally we can see the visualization of the

information in the fingerprint. So this is for operations teams to look into the characteristics of the attacks.

So these are two short demos of the components that we developed in this particular project.

Okay, so this is where we are from a technical perspective. Now of course, we want to test how the system works. So what we're doing is we set up what we call an anti-DDoS coalition in the Netherlands. So we call that a Dutch anti-DDoS coalition, and it consists of…actually, I'll talk about what it consists of in a minute.

But what this slide shows is that the software development and the development of the cookbook and all that stuff, that's being done in Concordia. And then the results are being used for the Dutch anti-DDoS coalition. So we really have some research and development going on in Concordia and then the exploitation of those results in the Dutch anti-DDoS coalition and specifically regarding the DDoS clearinghouse operations.

So the Dutch anti-DDoS coalition consists of 17 organizations. Some of them are partners in the Concordia project; others are not. And they consist of ISPs or Internet exchange points. Large ISPs in the Netherlands but also government organizations. SIDN so the .NL registry and a couple of other organizations. So it's really a cross sector. But of course, you can organize your anti-DDoS coalition for a specific sector if you want, such as the finance sector or the mobility sector or whatever you choose.

Currently we are in the state that the members have committed to freeing up a budget for this specific initiative. So for this year it's 114K. that doesn't sound like a whole lot, but stuff needs to be done because the system needs to become operational, software needs to be maintained even after the Concordia project. So this is actually a very strong signal that the coalition partners believe in the concept, and that's really important to actually get it into a production level because we're still in the research and development phase at this point.

So what we're also working on is I previously spoke about that this is not technology because it's not only technology because it's also about developing agreements to share information with each other. And that's important because under the GDPR IP addresses which are part of the fingerprint are considered PII. So that's something that needs to be taken care of through legal contracts. And then we also have a consortium agreement which is currently being fleshed out by a group of legal experts to create the real consortium, so to speak, as a single operating unit.

So our planning for the rest of the year and also for next year is the during Q2 and Q3 we want to continue with the development of the DDoS clearinghouse. And we would like to run a pilot with it so that at least a few of the Dutch anti-DDoS coalition partners connect to it. So we're aiming for three, and that's already quite challenging because people need to change their operational infrastructure in order to get the dissector deployed in their system, in their infrastructure, so that's really a challenge.

And then our next phases are to continue the development of the DDoS clearinghouse in the Concordia project but then make a fully operational service out of it. So in that case the test setup that we're currently running at SIDN Labs will go somewhere else and will be run by professional people doing professional hosting and all that sort of stuff.

And then in the Phase 2 we'll also phase out the software development. So the software development currently takes place in the Concordia project by these different partners that I mentioned previously. And at some point this will also need to move to a commercial software development company that can take over fixing the software and further improving it.

So it's basically three phases—pilot, basic production, and then full production. And that's also when the Concordia project ends which is going to be at the end of next year. So this is our happy path, so to speak.

I already mentioned that before but an important challenge is that the dissector can be run in partners' networks. And to be able to do that, they need to create some sort of mirror port of another type of splitter, if you will, where the DDoS traffic is being fed through the dissector and the dissector can create the fingerprint. So usually that's first doing a sample of the incoming DDoS traffic and then feeding it through the dissector. So what folks need to do is basically set up the branch at the bottom. So get a snapshot of the attack, feed it through the dissector, and then generate the fingerprint.

Another challenge here is that many organizations may have partnered with a commercial DDoS mitigation service. So when the DDoS traffic comes in they try to handle it locally initially, but when they cannot handle it anymore they forward the traffic to their mitigation provider. But that means that the dissector loses its traffic, so this is something that we also need to look into. So probably this will also require the collaboration with third-party DDoS mitigation providers.

This is our planning for the rest of the year. I think I already talked about that. Getting the DDoS clearinghouse to couple with production networks of the partners, maturing the clearinghouse's components, and then also making the results available through this open source design, if you will, in the form of a cookbook [inaudible] paper or something. So this is our outlook for the rest of this year.

I included a bunch of blogs here that you guys can have a look at if you're interested.

That sums up my presentation. Thanks. If there are any questions, I'd be happy to take them.

EBERHARD LISSE: Can you do the few slides, thank you, for the host presentation now or we want to do them later after the questions?

CRISTIAN HESSELMAN: Maybe after the questions. That would be more interesting.

EBERHARD LISSE: There is one from [inaudible] Q&A pod which you should also be able to read. Let me read it aloud. Can you explain how much money/admin power was invested so far into the system? Is it funded by Concordia [2020 EU]? This project has received funding from the European Union's Horizon 2020 Research and Innovation program under grant agreement and the number.

CRISTIAN HESSELMAN: Well, it's difficult to put a number on it. It's a very good question, by the way. We're getting funding from the European Commission to do part of the work. So for example, SIDN and the University of Twente and Zurich are being partly sponsored by the European Commission to do this work. But in the Dutch anti-DDoS coalition there are 14 other organizations and they don't get that funding. So they put in the money in kind, so to speak.

So I'm not sure what kind of number I should put on it, to be honest, but it's quite a bit of work. Not only from an engineering perspective and a research perspective but also for the working groups that I just talked about. So there's a legal working group. There's a working group looking into DDoS drills. So we also do DDoS drills in the Dutch anti-DDoS coalition. That also takes a lot of time to organize, but it's very valuable. And we also do information sharing, so sharing of expertise, if you will. So a lot of people are putting in a lot of time.

EBERHARD LISSE: There's another question from Syed Shah again. Do we ask that this whole ecosystem that you presented can be a key component of a national [CERT] of any state/country? Before you answer, you mention it's open source, so that probably answers it already.

CRISTIAN HESSELMAN: Yeah, so the idea behind making it open source is that anyone can benefit from the work. So not just us but anyone in the world. But we believe that the Internet is maybe the world's largest collaboration because there are 70,000 networks somehow interconnected. So if you want to protect that, then you probably need to work together a little bit. And we think that if you really want to do proactive DDoS mitigation, you cannot do that on your own. You need to collaborate with other organizations. And we think that this is one way of doing it.

There are other measures that you will need to take. So for example, if you look at the [edges] of the network, you will also need to make sure that your IoT devices are secure so that they cannot launch these Mirai types of attacks. So I think there are different measures to take through the Internet infrastructure, but I think that this certainly is a key component in that equation.

EBERHARD LISSE: I have always like collaboration, especially as far as open source is concerned rather than hiring commercial providers to do that. Anyway, thank you very much.

CRISTIAN HESSELMAN: You bet.

EBERHARD LISSE: Can you do us a few slides about SIDN, please, now?

CRISTIAN HESSELMAN: Yes, I can.

EBERHARD LISSE: We have just about six minutes' time, so we can do that.

CRISTIAN HESSELMAN: Oh, I have too many slides in that case, but I'll [inaudible].

EBERHARD LISSE: Oh, don't worry. Take your time. Ray has got more minutes than he needs, so you can exceed a little bit. It's no problem.

CRISTIAN HESSELMAN: Okay. Just let me look at the slides because I lost it somehow. Just a sec, Eberhard. I need to try it.

EBERHARD LISSE: No problem.

CRISTIAN HESSELMAN: I'm almost there. Okay, I got it. All right, close this one. I'm going to try to share it. Okay.

EBERHARD LISSE: There you go.

CRISTIAN HESSELMAN: Start presentation, here we go. Okay, so Eberhard asked me to do kind of a host presentation. We're not the host of this Tech Day, unfortunately, because originally it was supposed to take place in The Hague which is in the Netherlands. So that's why Eberhard kindly asked me to say a few words about the .NL and the registry role that we at SIDN fulfill for that particular country code top-level domain. So here we go.

So SIDN, we're the operator of the .NL TLD. That's what it says there. The Netherlands is a densely populated country, 17 million inhabitants, and we also have a dense domain name space, if you will. So we have 6.2 million domain names at this point, so that's quite a few domain names per person.

EBERHARD LISSE: Could you put the slide on full screen by pushing the play button, please?

CRISTIAN HESSELMAN: Oh, I did. At least I think I did. What about now? No.

EBERHARD LISSE: No. The PowerPoint has a play slideshow thing. There. No, that's the presenter. You probably have got two screens going.

CRISTIAN HESSELMAN: It worked just a minute ago. No. Okay, let me try again. Hold on. Stop sharing. Let's see if it works now. How about now?

EBERHARD LISSE: There you go. That's very good. Thank you.

CRISTIAN HESSELMAN: Okay, you bet. All right, so let's see if this still works. Okay, so we're the operator the .NL TLD, 6.2 million domain names at this point. We also have a lot of DNSSEC signed domain names and we get quite a few DNS queries on our infrastructure.

Our goal is to increase society's confidence in the Internet and the Internet infrastructure, and we basically do that through two different strategies, if you will. One is to provide secure and fault-tolerant registry services for .NL, both for the DNS as well as for the registration side. And we also have domain protection services that come with it.

And the other strategy is to increase—so that's more of our public strategy—to increase the value of the Internet in the Netherlands and elsewhere. And that's something we do in two different ways. One is to enable safe use or novel uses of the Internet the SIDN Fonds. That's a

fund that funds innovative projects. And we have IRMA which is an identity management solution that we're working on. And also we have SIDN Labs which is the team I'm responsible for, and our goal is to increase the security and trustworthiness of the Internet infrastructure.

This is the person that started SIDN. He's the godfather of the .NL top-level domain. He registered it with Jon Postel in 1986, and then he founded SIDN in 1996 because there were too many domain names at that point. I'm sure it's similar in other countries.

So the number of .NL domain names is at 6.2 million at this point. And as you can see, it went through quite a steep increase in the early 2000s. And the increase also came, as of 2010-ish, with the DNSSEC when we started incentivizing registrars to adopt DNSSEC signing. Which worked quite well because we're now at 55% of the .NL zone, so that's quite good and we're quite proud of that, to be honest.

If you want to have more statistics on what's going on in .NL, we have a separate site for that. It's called stats.sidnlabs.nl, and there's all kinds of statistics on .NL you can look at. And it's almost real-time, so it's pretty cool.

So .NL usage in the Netherlands, the .NL TLD is by far the most popular. So around 63.9% of the domain owner use .NL, and then we have 25% for .COM, and the rest is other.

This is what our infrastructure looks like at a very high level. We have our registration system on the right where registrars put in domain names. Then we have a signer for our resigning modules for adding

DNSSEC fingerprints. And then we distribute our zone files through four different types of infrastructure. One is CIRA and partly also run by ourselves. We have part of our infrastructure run by Netnod and part by RCodeZero so that's nic.at. And also we have part at ISC.

So it's important to have heterogeneity and we work with different types of models, you could say. So different software platforms, different hardware platforms, all that sort of stuff. I'm sure this is pretty common across the industry.

This is what NS2 looks like which is being operated by Netnod. So all kinds of nodes around the planet. I'm sure I'm not surprising anyone here.

This is the registration infrastructure. We have our registration services distributed across three data centers in the Netherlands, and they're being cross synchronized through fiber optic links. And we have a 99.96% availability roughly and full automatic failover. So if something happens, the other data center can take over.

Other security areas that we're active in, well, obviously system monitoring and patching. So secure software developments. We do our own infrastructure penetration testing. We also do large scale and collaborative DDoS mitigation drills. So that's part of the Dutch anti-DDoS coalition which I just spoke about. We have our own security operations center. We're ISO 27001 compliant. And we also do a lot in abuse mitigation just like the folk at EURid. So phishing, malware, fake web shops, what have you.

Something that we're currently working on is a more flexible DNS infrastructure. What we're currently experimenting with is having multiple virtual machines at cloud providers around the planet. We're using three at this point with Vultr, Packet, and Heficed. Basically, it complements the [as a] service DNS infrastructure that we have so that we get from CIRA, for example, and our own infrastructure—and this is sitting somewhere in the middle—and it enables us to easily play around with BGP catchments. Which is the map you're seeing on the right, for instance, by manipulating BGP path prepending or with BGP communities. So it's all about a more flexible DNS infrastructure.

By the way, the numbers between the right brackets on the top right, those refer to blogs which I have included at the end of the slide deck if you're interested.

We're also working on this system called I Reveal My Attributes (IRMA) which was developed by a university in the Netherlands. And that's something that we took over actually about a year ago. So we are now doing the software development for it. It's an open source system for decentralized identity management. So this means that your credentials, your personal attributes like where you live, what age you are, your bank account, all that sort of stuff resides on your mobile phone and does not reside in a profile somewhere with a service provider.

So this means that the users have the real control over the data because they decide when they release a certain type of information with a certain service provider. So that increases users' data autonomy. They

have more control over what data they share with whom. It also reduces the level of user profiling that large companies can do because the information is now on the user's side rather than on the company's side. And it also enables security verification because the whole system is open source. And it's already being used by two municipalities in the Netherlands and an insurance company and there's more to come.

So then SIDN Labs, that's the team I'm responsible for. Our goal is to increase the trustworthiness of the Internet infrastructure. So that's really the communications substrate that we all build upon. And we do that through three different strategies. One is through applied technical research. So that's doing large scale measurements, designing new systems or new prototypes, prototyping them, and then also evaluating them.

Like I said previously, we make our results publicly available because we think of ourselves as a private organization with a public role. So we try to help the Dutch and international Internet community with our results. So that means that everything that we do is open source. We share our publications and, if possible, we also share data.

The third strategy is that we collaborate a lot and we work a lot with universities, for example, but also with infrastructure operators and other research labs. The research topics that we work on are network security. So that's really that core systems of the Internet like the DNS, BGP, stuff like that. Then we have domain name and IoT security. And we do work on trusted future Internet infrastructure. I'll have a few words on that later on.

Am I still good to go, Eberhard? Or do you want me to…?

EBERHARD LISSE: A few minutes you've got left.

CRISTIAN HESSELMAN: Okay, I only have a few slides, so I should be all right. Okay, so here are a few example projects. I'm just going to highlight maybe two or three. The one on the top left is about measuring the time it takes to standardize a new DNSSEC signing algorithm to a large scale deployment. Which is interesting because it's important for the future safety of DNSSEC.

Bottom left is a logo detection system that we use to detect domain names that use logos in a malicious way, so for example, related to COVID. And we also do work on future Internet infrastructure. So we experiment with a system that's called SCION which is what they call a clean slate Internet architecture. We managed to port that to P4. P4 is a language for programming switches in hardware. And that's something that we recently put on a testbed.

We also do some conceptual work, at least conceptual until now, about developing new Internet security paradigms. And we recently applied for funding with a couple universities for that concept and we received funding from the Dutch science counsel for that, so €1.9 million. And that's money that goes to the universities to carry out that research. So that's really trying to bootstrap the community of researchers working

on future Internet in the Netherlands, and that's panning out quite well so far.

So SIDN Labs, we're not an academic organization. We're also not operations. We're somewhere in between. So that's why we work a lot with universities, but we also work a lot with operational teams such as SIDN's, for example. And they're really important to get information, to get new insights into our team, both the folks at the operational team as well as the more academic people. So we're really trying to bridge these two worlds.

These are a few examples of parties that we work with. Like I said, universities, international, national, but also the network operators like the Amsterdam Internet exchange and Surf Netherlands.

These are the folks who actually do all of that great work. We have one vacancy at this point. We also work a lot with students from universities, and sometimes they're very good and we hire them. So this is really great.

The work that we do is we help our operational team, we analyze large amounts of data, write open source software, write papers, that sort of thing. And we're also being helped by master's students who advance our work in very specific areas.

These are the blogs that I referred to. Whenever you see something in square brackets in the presentation, it refers to this slide. So have a look if you're interested. You can always ping me if you have any questions.

That was my last slide. I hope it wasn't too much for a host presentation.

EBERHARD LISSE: Excellent. I like that large nonprofit registries pour some money into research and into development, especially when it's open source. Thank you very much. There was one question about the presentations. All presentations are listed on the public schedule. It's in the chat, and they will be coming on the website so we don't have to spend any more time in it. Thank you very much, Cristian, again.

CRISTIAN HESSELMAN: You bet.

EBERHARD LISSE: And I now call on Ray Bellis to do his demonstration.

RAY BELLIS: Yes, hello. Can you hear me okay?

EBERHARD LISSE: Yes.

RAY BELLIS: Great. Thank you. Yes, this is Ray Bellis here. I'm the director of DNS operations at the ISC which means I'm responsible for F-Root of the root DNS system. Next slide, please.

Okay, so I'm here to talk about a root system visualizer that I've recently built. This has some similarity with a presentation you may have seen six years ago, so actually back at DNS-OARC where I [inaudible]. But basically what this is is an interactive map which I've built to support the ICANN RSSAC caucus's working group which is described as a tool to gather a local perspective of the root server system.

I should clarify that this interactive map is not actually the eventual output of that working group, the tool of the title. But this is built to inform the development of that and to help bring on some new ideas into that group. Next slide, please.

What it produces is a global map. We have all the usual pan/zoom type features, where each dot on the map represents the latency of the root system as seen from various vantage points. It can plot the nth fastest root [inaudible] letter which then gives you a metric of how many roots are within a particular latency performance figure or it can also look at just one individual root letter at a time.

The purpose of this is to help to be able to visually identify areas of the globe that are potentially underserved by the root system. Caveat that very carefully potentially underserved because it doesn't necessarily mean that the root server system itself is actually under provisioned in any area but may indicate some other issues with the routing and peering setup in different areas. Next slide, please.

The data source for all this is the RIPE Atlas system, which Eberhard mentioned earlier. It currently has in excess of 10,000 network monitoring devices that have been given out at various network

meetings, particularly RIPE but many others as well. The nodes are unsurprisingly concentrated in the RIPE area because that's where they've mostly been given out. They're typically the size of, depending on which version you have, a stick of gum or maybe slightly larger in some cases on newer versions. But they're very small and they're just given away by RIPE.

These are capable of performing lots of different sorts of measurements out to the Internet. The ones that I find of most use are they can ping any destination, can record traceroutes. But it can also perform DNS lookups and you can specify what particular resource record is to be looked up, whether it's IPv4, IPv6, etc.

There is a standard set of built-in measurements that RIPE built into every probe, and the results of those are freely available and are what's actually been used to generate this visualization. But you can also build your own custom measurements so long as they use the protocols that are supported by the RIPE system itself.

All of the data that RIPE collects is stored in a very large data processing cluster which I believe is still based on Hadoop, and they expose that via a REST API so you can look up current data and even actually go back and to historic data. In fact, even real-time mode, you can have a website based real-time drip feed off the results that come in as and when they're reported by the individual probes. Next slide, please.

So [inaudible] measurements, as I mentioned, there are automatic built-in measurements of the root system. And specifically they send a

hostname.bind CHAOS TXT query to every one of the 13 root letters. It's roughly every 10 minutes or so.

Because of the IDs that are built into the system, this also means that we can identify which site is being probed generally. The vast majority of the RSOs use hostname.bind names that are based on an IATA or airport code or an ISO LO-CODE. And we also, of course, get query latency from those.

And RIPE also happily provides us an API which tells us the latitude and longitude of every probe. So given all of those together it's possible to make a nice graph model. Next slide, please.

My slide's not [inaudible] the slide after. Could you go to the next slide, please. Yes. Ah, sorry. Kim, I think [inaudible] slides are in the wrong order. Sorry about that. Okay, this is supposed to be the next slide. That's good.

So this is the global view of what this looks like. The dots represent the latency as shown on the scale at the top right. Currently this view is showing the fastest root letter that's been observed by each probe. Because the dots are so many on the screen—so there are 10,000 points plotted here—the ones with the slowest latency are plotted nearer the front of the graph, as it were. So nodes with a poor, apparently, reported latency will always appear more prominent in the graph. Can you go…so, Kimberly, can you go to the slide please that says less than…at least one RSI? That's the one.

Okay, so a zoom-in here of western and central Europe. What we're actually seeing here on the scale is that actually the vast majority of probe locations—[I see one RSI] where there's less than 100-millisecond latency to any node on the root system, whether that's A through M.

The red dots in this case tend to be ones where they've got particularly slow local loops for whatever reason. We have not been able to dig into all of these and find out why that is, but it might be for example they might be using a satellite link or some other sort of very slow Internet connection method.

Do bear in mind that most Atlas probes are actually often located in the homes of techies that received those. So you're not necessarily seeing the latency of the root system as it would be seen from a recursive resolver sat inside a data center but often a couple of hops further out, maybe at the end of the DSL line. So that's those red ones there. Okay, next slide, please, Kimberly. The one that says at least three. I apologize for the confusion here. My slides got put into the wrong order. There we go. They got put into the wrong order in the PDF.

Okay, so this is the same as the previous graph of Europe we saw. At this point I'm actually coloring this by the third fastest root instance. So that basically means that I'm looking at, well, it's not the third percentile but 3 out of 13. So there are two root servers at least faster than the color that's been represented here. It's generally looking pretty good. You might be starting to see that some of the colors are getting a little bit less green, particularly in France in this case. I see more slightly off-green dots there than there are in other parts of Europe. Next slide,

please. The seventh fastest instance. It should now be Slide 9. There we go.

Now we've changed it to look at the seventh fastest instance. So again, a lot of France is looking quite yellow rather than green. Actually, [inaudible] the whole of Iberia, Portugal in particular, and eastern Europe is now starting to look somewhat worse.

But I should qualify that by saying this is based on only 100 millisecond latency. If you download and play with this tool yourself or just run it off my instance, you'll see at top right there's the slider where you can change what sort of color band you're looking for. On F-Root we're generally trying to get within 20, 30, 40 milliseconds [everywhere.]

But that notwithstanding latency across the planet can be quite poor sometimes, and it's not always deterministic necessarily which [site] you'll get to see. Next slide, please.

This is how things look if you're actually just looking at the slowest of the root letters across the whole of Europe. The takeaway here is not so much that there's lots of red dots, but the surprising thing is there are so many green dots, specifically in Germany, predominantly Germany. Essentially, what this graph is showing is that every single root instance has less than 100 milliseconds latency to all of those green dots. In some cases much less than 100 milliseconds latency. Germany appears to be spectacularly well connected. My guess is that a lot of those more yellow/orange dots in the eastern side of Europe are connected back into major German ISPs. Next slide, please.

This is a picture of what happens when you just look at one root letter. In this particular case it's G-Root run by part of the U.S. Department of Defense. This is a slightly curious one because you can see there's a very, very clear split here between France, Iberia, and the U.K. from all the rest of central and eastern Europe.

I've been digging into this in a little bit more detail and what I'm finding is that even those red dots [typically] go back to the only two locations that are run by G-Root in Europe. But those locations I think are Stuttgart and Milan from memory. And what I'm seeing is I believe is a question that these probes are having to back call all the way to a national—presumably probably a capital city data center and then go over to another capital city data center.

So for example—or even not a capital city, at least a major city data center. So you might be going from the south of France to Paris and then from there to Hamburg and from there to Stuttgart. So it's quite important to get good root latency that peering actually if possible it be expanded as far as possible out and distributed away from the national centers.

[In fact we] just deployed, put a root server in Malaga at their request for a new IXP that's been formed down there specifically because they wanted to avoid the fact that all the root system traffic is going all the way back to Madrid. Next slide, please.

As I was saying, seeing the poor latency on that graph does not necessarily mean there aren't enough root servers in an area, but it could actually just mean that the BGP mesh is not as good as it could

be in a particular area. So peering is vitally important. And if you're a long way from your capital, from your major interconnection point in the country but your quite near to a neighbor, actually seeing if you can get cross nation links to your nearer neighbors might actually get you better root service than you'd expect.

The other [inaudible] we've seen in the past but fortunately don't suffer from too much now is also don't try to peer with root servers over a long-haul link. Because we've often seen problems where you've got private [inaudible] and going internationally and they can cause quite a bit of problems because then you get root service…we've had an existing problem with the [inaudible] and the satellite two network where there's one point in the network where they were treating our F-Root prefixes as a customer prefix and sharing it with every single international academic network that they announce roots to. The net effect being that international research networks all the way from Japan, the U.S., whatever were all going backhauls to the west coast of the U.S. which is not optimal.

I don't now have time for a live demo, but I've even used this tool for even [micro tunings of our] connectivity. I'm just down the road from the Oxford University and I happened to spot a couple Atlas probes there on JANET's AS 786 which were inexplicably going to Charles de Gaulle. Sorry, it actually wasn't the [inaudible] it wasn't [inaudible]. It was Charles de Gaulle in Paris they were going back to. And AS 786 was going to Paris likewise.

We spotted that and resolved that by making sure we had a local peering in the U.K. with JANET. And now instead of—okay, it's small numbers but instead of 12 milliseconds latency to F-Root they've got 4 millisecond latency. It's a small step and it's a micro optimization, but it's something that we couldn't necessarily have seen otherwise. Next slide, please.

Okay, well, we didn't have time for that, so final slide, please. Okay, thank you. Any questions, please put it into the Q&A. I have a running instance of this. Anybody can access it at the URL shown here. The source code. It's a tiny little web app. It's developed from GitHub so anybody is welcome to download that and run it locally. Ideally Google Chrome because I don't have time to worry about trying to make sure this works on every single browser. Thank you.

EBERHARD LISSE: Okay, thank you very much. Great stuff. I had been playing with this a little bit, but I don't really understand the details of root server management. So it becomes much more clearer. There was a question in the Q&A pod from [Hank Nussbacher]. It doesn't work. I can see it. It does work for me, and I can see my both probes, my three probes in Namibia, two in Windhoek and one in Swakopmund.

One thing I wanted to remember if anybody wants to do an Atlas probe, you can download the software. I've put that into a little Raspberry which runs at my house as well as where one of the hardware probes is running and that works very well. Eventually I will have another

Raspberry on the same network as I have the hardware probe so that we can see whether hardware and software probe make a difference.

Anyway, thank you very much. Great stuff.

RAY BELLIS: Yeah, if Hank wants to reach out to me directly, we can try and find out why it's not working for him.

EBERHARD LISSE: It worked for me, and in Namibia our connectivity is not that good. So he'll figure it out.

RAY BELLIS: Yeah, it does take about…when you first access the site, it does take about a minute to download all the data from RIPE's API server. But it should show a progress bar while it's doing that.

EBERHARD LISSE: All right, thank you very much. We have a short break. We'll start exactly in 25 minutes, half past 2:00 UTC. Thank you very much.

UNIDENTIFIED MALE: Eberhard, there's one question from I assume Brett.

RAY BELLIS: Yeah, sorry. I was trying to reply to Brett. Brett, the code would work for any…yeah, actually, I hit the wrong button. Yeah, it's possible to modify

the code to work with any measurements that's exposed in RIPE Atlas. You just need to put in the right measurement numbers. So it wouldn't take much to tweak the code for that.

EBERHARD LISSE: Brett, if you do something like this, please email me so that we can see you present what you've done on the next or in a future Tech Day. All right, thank you again. Let's go and have a short break, and then we'll meet again at half past 2:00 as it says here on the slide.

[END OF TRANSCRIPTION]