BARCELONA – Tech Day (3 of 4)
Monday, October 22, 2018 – 13:30 to 15:00 CEST
ICANN63 | Barcelona, Spain

EBERHARD LISSE: Welcome to the [inaudible] session. Somebody is trying to threaten me with [inaudible]. Thank you very much for coming back, even though we are less than we left. Some attrition. Next presenter will be …

UNIDENTIFIED MALE: Bruce Tonkin.

EBERHARD LISSE: Next presenter will be Bruce Tonkin from auDA, who managed to convince the Guinness Book of Records that they hold the world records in the transfer of domain names, and he can tell us all about it.

BRUCE TONKIN: Thank you, Eberhard. So, let me see. So, I'll give a little back of a background as to what we were doing last year in terms of the process to select a new registry operator. Most of this presentation will talk about the transition and the technical steps that we undertook, and then we'll finalize with lessons learnt that might be useful for people that are also contemplating some form of transition.

So, I guess I'll start just with a little bit of background. The way – .au and the organization that I work for, auDA, is very small. It's only 15 people. We don't run the registry ourselves. We outsource it. It was originally

outsourced back in 2002. Each time we've outsourced it, it's a four-year contract. So, basically, it was a tender run in 2002. The Ausregistry in Australia was the successful operator. Then, they resigned three times, essentially. The last signing was for the period 2014 to 2018.

So, last year, we undertook an exercise going out to market again to select a registry operator. In December of last year, we selected Afilias to be the new registry operator, and, therefore, needed to manage a transition, which is, as Eberhard mentioned, the largest transition that has been done to date.

The previous slide's transition was .org when the Public Internet Registry transferred it from Verisign.

Our project goals. We're very focused on stability of the top-level name. A big focus in the tender round is looking for solutions that offered a high degree of security, and also gave us the opportunity to do some more data science and data analytics.

In the past, auDA ha completely outsourced the registry operator to the level that auDA that no data, either. So, we had no access to the data of the registry. Now we do.

To give a little bit of an idea of scale and, I guess, complexity for .au, .au currently has registrations at lower levels. They register within .com.au and .net.au and .org.au. So, we have just over three million registrations in total, and 2.8 million of those are in .com.au. Then, about 10% of the total amount is in .net.au. .org.au is for non-commercial. Then, we also have education and government. So, quite a few name spaces. We even

have one that dates back from the very early days: a single domain name in cont.au, which is the Linux conference in Australia that has been running in Australia since the late '90s. So, they have their own domain name.

We ran, as part of the tender, an expression of interest process, which we started in May of last year. We got 15 responses, a mixture of responses from ccTLD and gTLD operators, and also parties that wanted to develop the software themselves. So, we ended up really getting responses from a large cross-section from the major cc's and gTLD operators.

Based on that level of response, we thought it was worthwhile proceeding to a formal tender, which we issued in September of last year. We got nine complete responses to that tender. We shortlisted three and asked for best and final offers, which happened at the Abu Dhabi meeting. Then, we negotiated with contracted with all three of them and made a final decision in December.

So, once that decision was made, then we sort of had to rapidly get into the planning. The general approach we took is to try and align with international standards for risk management. For security, it was the ISO 27000. The Business Continuity – so, disaster recovery. We referred to the ISO 22301 standard. Then, for service management, you can use either ITIL or ISO 20000. So, there was some Australian security standards that we also adhered to as part of the process.

So, getting into the transition itself, we had six months from when the Board made its decision in late December. Everyone pretty much, in

Australia, ran away for Christmas holidays and then came back and started to work on the transition. The target was to transition by the first of July of this year, which we achieved.

But, we had a lot to do in six months. So, firstly, the Afilias software had to be customized for the .au requirement. So, we had customized data fields in the registry, and we have customized EPP extensions. So, it's quite a bit of software development that needed to be done on the Afilias platform to meet our requirements.

We had a large data migration exercise of three million domains, but associated with each of those domain names were numerous hosts and contact records. So, really, you're looking at more like ten million records that needed to be transferred.

For the infrastructure, our requirements were that the infrastructure had to be built in Australia. So, that included the registry platform itself, the disaster recovery [site]. We also decided to put DNS nodes in each Australian capital city.

One of the things that's is interesting is that, in Australia, we've got a small population that's about 25 million people, but most of the population lives in quite large cities. So, each state or territory of Australia has its own capital. Most of those capitals have more than a million people in them. So, they're fairly large cities, and we decided we wanted to put DNS infrastructure in each of those cities.

Then, we tried to ensure a smooth a transition as possible with registrars, so we had monthly meetings with registrars.

We took a risk management approach to the transition, so we used the ISO risk management standard, and then we also appointed an independent risk auditor, which was Ernst & Young. They were providing independent advice to the Board of auDA on the transition process. That advice was also shared with the Australian government.

On the software development side, we ended up doing two drops of software. The first one was in March. So, we started in January. In March or, really, late February, we made a test environment of [inaudible] for registrars. So, that basically gave them four months, really, to test the new platform and make sure their systems were working.

The first test environment included all the core operations to manage domain names. So, the standard EPP – create, update, renew, cancel operations.

Then, the second test environment included more specialized commands. So, we have quite a few specialized commands. In .au, one of the things we do is we have a strong compliance approach to enforcing the rules around the .au policy. Then, we have different types of delete operators. So, a standard delete operation means that, after the domain is deleted, it becomes available for re-registration in three days. But, we also have the concept of a policy delete, where, if we think someone has infringed the policy, the name gets deleted, but they actually have 14 days to appeal it, basically, before it gets purged from the registry and made available for resale. So, that's a separate operation.

We also track transfers between registrants. So, we have a separate command. It's not just a standard update command, but a separate command, so we can track sales, basically. So, we can see when there's domain name sales going on in the market per the registry platform. There's a few other commands as well.

So, the second test environment was deployed in April, and it was fairly stable. There were these odd, small, I suppose, bug fixes and things along the way, but those were the main deployments. So, we aimed to get as much testing of the system as possible.

We essentially had three layers of testing. So, Afilias its own testing. In auDA itself, we engaged people that had backgrounds in the registry and registrar business to do testing, particularly of the EPP command environment. So, we wanted to make sure that we independently had assurance that the registry was meeting the requirements that we put in the technical specification. Then, registrars did their own testing.

What we found with registrar testing – I think a lot of gTLD operators are probably seeing this – is that registrars tend to do a minimal amount of testing early on, really just to explore what's new in the environment. Then, they tend to do most of their testing within a couple of weeks of the deadline, when they suddenly panic and realize, "Hey, we've got a transition in a couple of weeks' time. We better start testing for real now." So, part of that was why we wanted to do our own testing in advance.

We picked up a few things along the way so that, by the time registrars were testing heavily, there really was no issues that they found. So, the registrar process was very smooth.

On the data side, we gained access to the data from the previous registry operator in March, and we would get a monthly – bear in mind, we didn't have direct access to the data, so we had to get the data from the previous operator. So, for March, April, and May, we got a monthly drop of the data. So, a complete set of their data across those three months. Then, when we went into June, we got a weekly update on the data. Then, obviously, on the transition day, we got a fresh update of the data.

A lot of work then goes into transforming that data. The data format that received was an intermediate format, so, all the schema and data structure that was in the previous registry was stripped out. It was just like a raw dump of data. The data is not linked, so it's just, "Here's all the domain information. Here's every host record information. Here's every contact information." Then, you had to kind of rebuild a database of that data. So, that was a big exercise that Afilias had to do.

One of the things that we needed to do is to try and validate that the migration was accurate. So, the first thing we used there was to test the zone files. So, we're able to have read access to the zone files. So, when we got each drop of data, there was a conversion process that converted that data into the new registry database.

From that new registry database, we generated a zone file, and then we compared that zone file with the zone file from the current operator, just to ensure that all the domain name information was accurate.

The other thing we did was we created a publicly-accessible WHOIS service that was generated by Afilias from the data that it had received from the previous registry operator. That meant that registrars could audit and check that domain names that they knew that were on their platform and could actually order the WHOIS and say, yup, the data that they had, if you like, put into the registry was coming out of the Afilias registry in the right format.

So, a few things got picked up there, like some slight differences in statuses. So, there were some statuses that weren't provided by the previous operator but which you could see in the previous operator's WHOIS. So, when Afilias generated their own WHOIS from the data that they had received, it was possible to see [inaudible] registrars flagged in a few cases that there was some missing statuses. Then, we were able to go back to the previous registry operator and ask for that additional information.

So, basically, [with] every single one of these drops, the data changed slightly, and we were able to potentially ask for, as we identified missing data, that missing data to be provided so that, by the time that we got the transition date itself, we knew that we had data integrity from the previous operator to the new registry operator.

Certainly, if you're doing a major migration, this is probably the critical thing to get right: you get the data migration right, and you constantly

test that and start the testing as many months before transition as you can.

The infrastructure rollout. In parallel with all this, Afilias was building infrastructure in Australia. We stipulated that the registry data couldn't leave the country. So, for any testing they did, they had to build test environments in-country.

Before we would let them put any registry data onto a test environment, we required the testing environment to go through a security check. So, that basically meant that Afilias had to provide auDA with a copy of the security order that they had done on that environment, as well as independent penetration testing of that environment. So, we basically went through that audit information before we'd allow them to move data, even into a test environment.

We also did full disaster recovery testing before we would allow them to transition. So, they had to prove to us that they had transitioned the registry from one city to another.

So, in Afilias's case, the primary city is Melbourne, where their primary registry is. Their backup registry is in Sydney. So, they had to validate that they could transfer the registry from Melbourne to Sydney before we would allow them to transition from the previous operator.

The other thing we did was run Afilias name servers in parallel with the Neustar name servers. So, at the top level, the .au level, which has only got less than 100 entries – it's got .com.au, .net.au, etc. – from January of the year, we started running Afilias name servers at the top level in

parallel with the Neustar name servers. So, we essentially had six months of parallel operation between the new registries' operator name servers and the previous ones.

Then, at the same time, Afilias was adding name servers in each of the capital cities, so we were progressively adding DNS name servers over the six-month period between January and June.

So, the general approach we had was to try and do as much in parallel as possible so that, when we got to transition day itself, we already had months of operation of the new registry operator essentially in production. So, this was to minimize the risk in the transition itself.

The DNS transition. So, at the top level, we were running the Afilias name servers live in the .au zone. At the second level, we were running copies of the zones. So, .com.au was actually across to the live Afilias name servers, which we could query by querying the Afilias name servers directly. But, those name servers weren't published in the zone, so a consumer wouldn't naturally hit one of those name servers but were able to check that the name servers were operating and providing the same information as the previous registry operator one.

One of the things that I guess was perhaps a little less common is that we're also transitioning DNSSEC. So, the previous operator was signing the zones and providing a DNSSEC registry. The new registry operators, obviously, chose to use their own DNSSEC signatures independent of the previous registry operator.

So, we had to transition the DNS signing operation. The way we did that is, when the new signatures were actually published into the zone from Afilias in parallel with the Neustar signatures – so, for some weeks, both signatures were actually in the zone file and could be cached and accessed by resolvers. All we did on the transition day is changed from Neustar signing to Afilias signing. The signatures had already been cached and exposed in the zones.

So, that meant that the DNS transition was pretty much 99% complete before we even got to our transition weekend. Really, the only change by that stage was the signing and the publishing of the – the authoritative publisher of the zone was now Afilias instead of Neustar, but the DNS name servers had been operating in the zone for months.

We turned on the Afilias name servers at the second level in June so that, even for consumer queries, they were starting to hit the Afilias name servers. But, they were getting the zone file from Neustar. So, we made that DNS transition to be very seamless.

We had a lot of focus from the Australian government during the transition. Because of the size of the transition, they were concerned.

It was interesting. In Australia, the context was that, about a year previously, they decided to do a census check. This is where everybody in Australia has to fill in a form that says where they lived in the last several months and are asked a whole bunch of, essentially, personal questions. Last year, they decided to that do online rather than physically mailing pieces of paper around. They day that they turned the servers on, the 25 million Australians then went to submit their

census, and it completely crashed the census database, which was massively embarrassing to the government people that were meant to provide that service.

So, they were paranoid that the same thing was going to happen in the .au transition in that we would transition from one registry to another and all of .au would go dark and they would lose government.

So, we had a massive amount of focus. So, we engaged our own independent advisers, which is Ernst & Young, but then the government then appointed, separately, a consulting company to do its own review of our plan and process for transition.

Then, they set all the government security agencies on us, so we spent weeks answering questions from everybody you could image. So, we got through that. Then, we also had oversight from auDA's own Board's Security and Risk Committee.

In fact, we have two people that I can't see because the room's so big. But, we have Joe Manariti and James Deck, who are on the Board's Security and Risk Committee. So, they kept a pretty close eye on us as well. So, I think, was we got closer to the transition, we were meeting with the Board's risk committee monthly and meeting with every government agency daily.

Finally, the day before transition, the government said, "Yeah. We think we've asked as many questions as we possibly can, and we think we're comfortable with the transition." So, we went ahead.

Another key part of the process from the government angle was that they wanted to be clear that we had a crisis management plan. This is actually a good thing to have in place anyway. So, crisis management really is not just the technical side of it. Obviously, if something breaks in a registry at any time, you need a process for dealing with a technical problem.

But, when it's of a large scale, you're also needing to look at how do you keep all the stakeholders informed in the middle of a crisis. So, the scenario that we were looking at is: the .au registry is relatively large. It's sort of three million names. We have about two million businesses in Australia. So, yeah, probably one-and-half million of those businesses have domain names.

So, if .au went down, it would have a serious impact on the economy, and we would be getting questions from the government. We'd be getting questions from the media. We'd be getting questions from registrars. So, it's, "Hey, you actually manage a crisis where you're dealing with not just the technical side but almost being DDoSed by people ringing you up and wanting to know what's going on." So, we got some professional advice on how to do crisis management.

Then, they key with crisis management is you've got to practice it. So, we actually created some scenarios. One of the scenarios that the government was concerned about was that, because of the fact that we were doing a transition that was public and we were publicly announcing that the date that we were going to do the transition was

the first of July, we could be a target for a concerted attack on the infrastructure during that time.

I think the government at the time was concerned that, in a moment of crisis, you'd attack .au, and then someone would send in some boats and steal all our beer and wine while were so focused on dealing with the .au crisis.

So, we ran a number of scenarios, where we had scenarios we had a DDoS attack from a hostile foreign power. We engaged not just the Afilias and auDA teams but the Australian government security agencies – these were people from the sort of Department of Defense in Australia – and just worked through a scenario of what would we do if we were attacked by a foreign power that was trying to destroy e-commerce infrastructure. So, we ran that scenario.

Another scenario we were running to was an attack through staff within the organization. So, this is the concept that someone has kidnapped the children of the staff member and threatened to kill them if the staff member doesn't destroy .au somehow. So, we're running through scenarios of, what if a privileged person who has got a lot of access to systems somehow gets compromised, possibly under a threat to their family? What damage could they do? How would we manage that? And, again, how would we manage a crisis around what was referred to as an insider attack.

So, we ran a few crisis exercises, and we'll continue to do that as ongoing process. That's part of our business continuity approach. It's not just to have something written own on paper but actually run live

tests. So, we run tests of a fire happening in the office and nobody being able to access the office, we run a scenario of a staff member that's gone rogue and is trying to do something malicious to the registry, or we regularly plan our process in defense, I suppose, for DDoS. So, crisis management was a key part of the planning for transition.

It was funny. Within the people that we're working with, at least for me, this felt very much like planning for the year 2000, because I remember being in the IT industry at the time. There was so much concern that, when the clock ticked over to the year 2000, hospitals would stop working and the world would come to an end. So, there was a massive amount of preparation. But, when the day came, nothing really happened.

That was our situation. The actual transition weekend, because we done so much planning and testing, was really smooth. Basically, nothing went wrong.

So, during the cutover itself, we planned to do the transition in 24 hours, but we announced a 48-hours window because we wanted to buffer in case something went wrong and we needed time to fix it.

We got very good cooperation from the previous operator, Neustar. So, they were exemplary in their support for the transition. So, because the data set was pretty large, it actually took them eight hours, for example, to just extract the data from their systems.

So, at the beginning, on the Saturday morning, we stopped accepted new registrations and then just basically had to sit and wait to receive the file of the data. Then, we needed to do testing against that file.

In the end, that occurred ahead of schedule. So, we were able to get the file from the previous operator and had plenty of time to analyze and test it. We were pretty much ready within probably sixteen hours or so. I think we had more or less done all the work for the transition. Then, between sixteen and 24 hours, it was really just really running a lot of tests. Again, auDA did its own testing separately from Afilias so that we could assure that the registry was operating successfully.

Then, we announced to registrars that the transition was completed around about the 24 hours. But, we more or less had completed the transition in about 16 hours.

Most registrars were live on the first of July, which is the first day of operation. So, they had all pretty much done all their testing. All they really needed to do was point the IP address to the new registry, and their systems pretty much came up smoothly.

So, really, from a consumer point of view, nobody noticed. That was really what we're trying to achieve: nobody in Australia would even notice that a transition occurred. And, that was mostly the case.

We also set up a network of probes, about 20 probes. We've got a probe in each Australian city, so we've got eight probes in Australia. Then, we have probes set up in data centers around the world. That's where we

measure the performance and the uptime of the registry from all of these locations.

So, this was new as well. auDA had always relied on the previous registry operator to self-report their own performance. We wanted to make sure that we had independent assessment that the registry was operating properly. So, we basically got our own monitoring system, and we're able to validate that the registry is meeting all its performance requirements when it went live.

So, the outcomes for us of the transition. We did get a lower wholesale price out of the transition, so we passed on a 10% drop to registrants. We maintained control over our data. It remains in Australia. We increased the number of DNS servers in Australian capital cities.

What tends to happen in Australia is that the biggest cities, Melbourne and Sydney, both have a population around the five- to six-million mark, so most IT companies, when they come up to Australia, put resources in Melbourne and Sydney. But, for the first time, we're able to put resourced in the more regional areas. Each of these cities is actually spread [out] by thousands of kilometers, so they are quite geographically diverse.

We also have proactive security monitoring now, so we're looking at names every day and how they're being used. We do a lot of data collection. We're still working at what we're going to do with all that data, but, as part of the new operation, auDA itself now has direct access to all the data.

Our focus is basically strengthening and go further with improving the security and performance.

So, lessons learnt. These are a few things that we have done differently, I guess. One thing that we found is that the EPP testing we did was very thorough, so that was where you'd send a command to the registry and then you'd read back the answer directly from the registry.

We had a WHOIS service that we set up with Afilias, but that WHOIS service was holding the current live data from the previous registry, basically, because we used the WHOIS service to test out of migration.

But, we didn't have a WHOIS display plugged into the test system, so you couldn't register a domain name in the test system and then see what it looked like when it came out through WHOIS. So, we did see some changes in the WHOIS display after the transition, and we need to make a few changes to how we were displaying WHOIS data.

The other thing is that, the earlier you get access to the registry data, the better. We only got access just slightly more than three months outside of the transition, so we got it mid-March. Really, we would have preferred to have got the data in January, which would have given us more time to help troubleshoot data errors.

Another thing that emerged that the registry operators had slightly different policies for posting host information in the DNS; so, IP address associated with name server records. In a nutshell, Neustar had a very narrow publishing approach. So, the approach was the publish the bare minimum and host data as necessary. Afilias had a wide publishing

approach, which was to approach any sort of IP address and host information they had they did make available in the DNS as additional information. So, when you're doing a DNS query, you have authoritative information and additional information, and Afilias published information in additional information.

So, it's kind of different approaches, I guess. There's not necessarily a right or wrong answer to this. They're obviously both running large registries. But, there was a difference, and, because of that difference, there were some old host records that got published that hadn't been updated for some time. If the authoritative name servers filed, then some of the DNS software would actually go and look for what other additional information was available. If that information was wrong, then you get the bad result.

So, we only detected about three cases of this actually being visible to a user, but it is something that, having known about that, we would test a little bit more, not just to see in the zone file what are the domain name records but just test a little bit more in terms of glue records and additional information and just seeing how it varies from the current operator to the new operator.

Another thing we discovered was that, in addition to registrars, in Australia we have many resellers. So, we only have a few registrars. We only have about 40 registrars. So, most IT service providers are resellers of a registrar.

But, the previous registry operator built some tools to make it a bit easier for resellers to manage transfers. So, there's a tool that allowed

you to retrieve the auth info in EPP directly from the registry, even if it wasn't a registrar. This was a tool that was heavily relied on by IT service providers that were web developers and hosting operators. When they got a new customer, their general process is that they would probably have a preferred registrar that they liked to use. So, when they get a new customer for their web design or hosting business, they would basically ask the customer to transfer their domain to their preferred registrar.

Because they didn't have direct relationships with the old registrar – and there's lots of registrars – they found it easier it easier to have an automated interface to the registry to get the auth info command or password.

So, we didn't have that particularly well-automated. In the first month, that was most of the customer service queries the registry was getting: IT resellers struggling to get auth info and passwords to transfer names from one registrar to another.

So, that has since got automated, but I guess it highlighted that, if we were doing the transition again, we'd put all our focus on registrars' monthly meetings. But, given the environment in Australia, where we have many thousands of resellers, we also needed to build a better communication approach, with resellers understanding what the issues that they might have in a transition are and making sure that we accounted for their needs as well as registrars.

Another thing that we found with the DNS name servers is we focused on geographic diversity and put the name servers in lots of locations.

But, particularly for IP version 6, we got some very strange routing, because, often, these name servers were being put in locations that really didn't have much use or heavy use of IPv6. So, the local telcos and IT service providers really hadn't set IPv6 up properly. Then, when we exposed these name servers and exposed them as being available by IPv6, we got some really strange routes.

So, [partly due to] the fact that we had monitoring probes everywhere – we had monitoring probes in, say, Amsterdam, for example, or London – we could do queries against each one of the servers from London and see where they were going.

It was weird. I think, at one point, Darwin, which is in a fairly remote part of Australia, seemed to be getting all the IPv6 traffic from Europe, even though there were name servers in Europe. So, Afilias needed to spend quite a bit of time working with service providers to optimize the rounds.

And, I guess that's something I'm sure many people [hear] across, but, yeah, it's not just physical location of servers, but it's getting the routing set up properly to take advantage of those so that the name server queries go to the closest Anycast node.

So, that's all I had Eberhard. I'm happy for any questions or anything else people would like to talk about.

EBERHARD LISSE: All right. Thank you very much. We have got enough time for questions. The first thing – I'm going to abuse my position from the chair a little bit

– is that , with this geographic routing, we see these two. We've got some in Namibia, a very small-place ISP, so they don't like to talk to each other. So, if you want to access a website from one part of time, it goes to South Africa, over [inaudible] to England, comes back, goes to South Africa, and then comes to the other side.

If you set up peering for these name servers in each location, so that all providers that are there can connect directly there, would that have made a difference?

BRUCE TONKIN: Yeah, it is. It's pretty straightforward in Melbourne and Sydney, but it's like what you say in these more regional locations. There's not the cooperation there, partly to do with cost and stuff, but the IT services providers is exactly the same thing. If you're in Darwin and you move from one service provider to another, you actually get sent via Sydney. So, we're trying to work to get that to work, essentially.

EBERHARD LISSE: [Ondrej]?

[ONDREJ FILIP]: I have a very similar question, actually. Thank you for the presentation, and congratulations to [inaudible]. It's a great job.

The question is, in the presentation, you mentioned that Afilias –

EBERHARD LISSE: Can you speak louder, please?

[ONDREJ FILIP]: Yeah – that actually has built service in each Australian capital, in eight cities. Did this cover the registry, or it was just DNS-related?

BRUCE TONKIN: DNS-related, yes. So, the registry itself is located in Melbourne. So, it's a central place for doing EPP queries. Then, there's a disaster recovery site in Sydney. But, we put DNS name servers in each of the capital cities.

I don't mean to say that that's the only name servers. So, Afilias has got name servers globally. But, these were name servers in new locations, basically.

[ONDREJ FILIP]: And those run as an Anycast [cloud]?

BRUCE TONKIN: Yes, they're Anycast nodes. That's right. The other challenge that we had in some of the locations was getting IPv6 connectivity. So, again, in Melbourne and Sydney, it was easier to do v4 and v6, but in some of the regional locations, we don't have v6.

What we have learned, though, looking at traffic monitoring – we're happy to start sharing some of this data with other TLD operators – is that we are getting a much higher proportion of IPv6 queries than I had

expected. I think we're getting about 20% of our DNS traffic as IPv6 now. So, that's dissimilar to when I talk to some other gTLD operators; they're not seeing anywhere near that higher percentage. So, it just shows you do need to cater for IPv6.

EBERHARD LISSE: Any other questions? Yes, come to the [microphone] and identify yourself for the remote audience, please.

[JOHAN IHREN]: Thank you for the presentation and sharing your experience. This is [Johan Ihren] from .pk. You have this thing about separation between policy and operations. Since the same constituents will be interested in policy and operations, how has the new structure worked, and how do you ensure that there's no conflicts? And, any best practices that you might share and how you put them together?

BRUCE TONKIN: Yeah. I think that the concept there between the separation of policy and operations is that we've generally outsourced the registry operations to a third party, but the third party doesn't have any control over the names policy.

So, auDA itself manages a policy development process for new policies. In fact, we're currently running a process to look at registration directly at the top level at .au, whereas currently you can only do it in .com.au and .net.au, and so on.

So, that policy process is independent of the registry operator. We invite anybody in Australia that has a view on that to participate in the policy process. We run public forums. We run public forums in several of the cities in Australia on that policy development process.

Once the policy is decided, then we would move over to operations and work out what's the most efficient way to implement that.

As you said, the stakeholders of implementation can be a bit different, so the registrars obviously, because they have to provide the service to the end user, will have a strong interest in the operations side of it, to make sure that works smoothly, as will IT service providers because, in many cases, it's the web developers. If we introduce a new name space, like .au, then those web developers will need to be able to advise their customers and help their customers with that.

So, the policy side tends to be very board. We get as many interested parties as possible, and we manage that at the auDA level. Then, on the technical side, which is mostly what we did in transition, we have stakeholders that we work with to get the best technical implementation.

EBERHARD LISSE: Any other questions?

Thank you very much. The next one is the Juan Antonio Gutierrez Gil will make it.

JUAN GUTIERREZ: Hello. Thank you very much for attending this presentation. My name is Juan Gutierrez, and I'm the Deputy Director of red.es [inaudible] the public corporate entity in [inaudible] the registry, .es in Spain.

We belong to the Ministry of Economy and Business, and, also, we developed a program in order to stimulate public programs regarding the transformation and regarding the digital economy.

In the next 15 minutes, I would like to share with all of you our trip to the cloud, a trip that we started five years ago and are finishing nowadays. This trip has been quite complicated, but I think that we are finishing in two months.

Five years ago, our service model delivery was based on technical assistance. So, if you need, for example, a specialist regarding a database, or a specialist regarding a database system, you just have to find this kind of specialist, or a company providing this kind of specialist. The [finished one] was very easy to launch, but it was very complicated to manage.

So, we started to think to change this model in order to change to another model based on services. So, we started a public tender procedure in order to sign just a single contract for a [uninterrupted] outsourcing.

Okay. After three years, last year, we found that [had] a single contract, and we had just one provider. But, this contract included the resale of our hardware assets. So, we haven't any [of our] assets.

So, this was the perfect situation to migrate to the cloud. So, we started a new procedure, a new tender, in order to migrate to the cloud. After one year, we have found this contract, and we're finishing right now. The scope of this contract is foreseen for four years. The budget is ten million euros. We have four transformation plans.

The first of them is regarding data center security and communications. The second one is related to work stations and end user services. The third one will be about corporate barriers; meeting rooms, auditoriums, and co-working spaces. The last one is regarding IT governance [inaudible] because we are too much people in order to match this.

Our environment is very complex. We have the [inaudible]platform, payment systems, open data, different webs, front ends and back ends. And, we even have our e-procurement systems. So it's a quite complex environment.

For us, we have four main objectives. The first of them is we must offer good responses to the business. We could even be providers of the other Spanish government departments. So, for us, it's important to be very flexible.

And, we have to act with efficiency because we need to establish a good capacity plan. [Co]-Computing is not always a cheap solution. You can't demand resources for beyond what you need. So, sometimes you can even not be aware of the final cost, of the final bill.

We must establish a different benchmark. We have to analyze trends. We have to analyze procedures. You have to analyze policies. We have

to analyze if we are compliant with the best practices in the market – for example, ITIL – and different international standards.

We must [inaudible] a contract where we have to be able to [inaudible] according to business needs. I'm paying just for that we use.

Here you will find our key topics. Our [licenses are mixed]. We have licenses. We provide licenses with this contract. Our [inaudible]. We have a part that is private, with Equinix as our provider. And, we have a public cloud with Amazon web services. Our [inaudible] are based mainly in our infrastructure as a service model but are flexible to evolution to our platform as a service or [through other service.]

Our service is based on [inaudible] [arguments], and we are compliant with our national security scheme and different international standards. We even have another contract with a service operation center that we have been integrated with. We are [complacent with] every regulation – European regulations, our data positioning in Spain, our internal regulations.

Here you can see our infrastructure. We have two data centers in Madrid. Then, we have an Oracle cloud, where we have our databases, and our public cloud is Amazon's web services platform.

For example, in [inaudible] service, we have Office 365. We [inaudible] all of our cloud with the conference systems, regarding streaming and [the things]. Our content management systems, based in [inaudible] workplace, are [inaudible] service with Oracle. And, our web front and

backend publications are in our infrastructure as a service. So, we have all computing models.

Here you can see our service portfolio or our service catalog. As I said before, we have three transformation plans. The first of them is regarding infrastructure communication's technologies. The three computing models: housing, software licensing, and communication lines.

Then, we have, all regarding our spaces, our office management, data processing, local networks – everything about corporate spaces. At the end is the end user workstations. Our workstation is mobile, so we can make work at home and these things.

Everything is [crossed] by an IT governance and politics based in international standards [inaudible], like ITIL.

For us, we are not a very big department. We are just three people and [inaudible] our things. So, it's very difficult to manage these. So, for us, it's very important to have a control panel, where you can see everything. How is the service?

We have a control panel where you can manage your computer resources. You can establish your service portfolio. We have [inaudible]. We have a [inaudible] called [inaudible] where we can monitor user experience. We have a platform called [inaudible] in order to see our resource consumption, business indicators, all KPIs, [inaudible] service level elements.

I would like to show you which will be our [inaudible] in the future. I think we are migrating in this one-month-and-a-half. It's a real complex platform, but I think that we are one of the first in Spain public administrations that is migrating to the cloud. I think we are the first.

So, here you can find our DNS systems with the [inaudible]. We have our security solutions and the denial of services or [inaudible] applications at this back end. These would be the private cloud.

Then, in Oracle cloud infrastructure, we have all of our [contacts]. For example, the [inaudible] registry, final user, different jobs. And we have, at the front end, [inaudible] web services. We have [dedicative] lines for everything in that. The main cities are in London and in Darwin, so it's a European framework. We are providing this, I hope, as I told you, I think in one month-and-a-half.

If you have any questions or any technical questions, you can write to me at this e-mail address. Here you can see different pictures of how we manage this infrastructure. We have [inaudible], online infrastructure, and control [tools] to deploy it in the cloud. We have different orchestrators. This will be [inaudible] cloud platform. We have the Amazon web service orchestrator. And, this would be the Amazon web service [inaudible].

Here you can see the Oracle cloud infrastructure orchestrator.

So, it's a very complex platform, but we are happy because I think that we could be an organization that could provide services to the rest of the Spanish government departments.

So, thank you very much. Enjoy your [inaudible] here in Barcelona, because the weather is fantastic right now. Now, I think this week is playing one of the best things in the world for Barcelona and for Madrid [inaudible]. So, I think it's –

EBERHARD LISSE: I got tickets.

JUAN GUTIERREZ: Really? For two days [inaudible]?

EBERHARD LISSE: On Sunday.

JUAN GUTIERREZ: Oh, Sunday. [inaudible] Barcelona. Good performance, I hope. Thank you very much.

EBERHARD LISSE: [inaudible]. Anyway, thank you very much. Are there any questions? How many registrations, how many registrants, do you have?

JUAN GUTIERREZ: 100.

EBERHARD LISSE: Now, how many domain names?

JUAN GUTIERREZ: Ah, domains. We have close to two million. I think one million and –

EBERHARD LISSE: And, your registrars use EPP, or what is it?

JUAN GUTIERREZ: Yes.

EBERHARD LISSE: Okay. With only three people. Very efficient.

JUAN GUTIERREZ: We've got in operations, yes, just three people.

EBERHARD LISSE: Okay. Thank you very much. Francisco Arias is over there – thank you very much. He will now to talk to us about another round of RDAP updates.

FRANCISCO ARIAS: Thank you, all. Hello, everyone. This is Francisco Arias from the ICANN organization. So, I'm here to talk about the RDAP implementation in the gTLD space and a status update. I think I talked to you about this topic in Abu Dhabi, maybe – something like that.

So, this is where we are now. This is the agenda for today, just a little bit of an introduction, the implementation status, and the next steps on the process.

So, first, how did we get here? We started almost nine years ago, back in 2010, I believe, or something like that. We started talking for the nth time, trying to talk with the community replace WHOIS. I'm not going to go through this list of issues. I think everyone knows the limitations of WHOIS. This is just a short list of the high-level issues that we have with the current WHOIS protocol, which is very limited.

So, on the context of ICANN, things started to get interesting in 2011, when SSAC published a recommendation to the ICANN community to [inaudible] replacement for WHOIS.

Then, shortly after the Board accepted the recommendation and directed staff to start working the community on this, we published a road map shortly after that, in 2012. Then, in parallel, the work started in the IETF to develop a new protocol to replace WHOIS. This took about three years. In 2015, the RFCs were published.

Since then, some TLDs [inaudible], particularly on the ccTLD space. I believe there are probably four or five ccTLDs that have already implemented this service, and they are listed in the IANA registry. I see Brazil here, and they are there in the IANA list. I believe the Czech Republic, Costa Rica, and Argentina are the ones that come to mind. So, there are a few that have started implementing this in production.

But, going back to the gTLD space, one of the things that identified that we needed to do with the gTLDs is that there was the idea that we needed to have something that we call a profile – the gTLD RDAP profile – that will define what are the things that need to be implemented from RDAP.

RDAP, the protocol, you can think of as a menu of things. It tells you how to do each of these things, but it doesn't tell you which things you do. That's a policy decision – what to do. In the case of the gTLD space, well, the policy gets made, usually, in ICANN, except for some things that are up to the registry.

But, in this case, the idea was to have a common profile that would be implemented across all the gTLD registries and registrars. So, the work on that started in June 2015. One year later, the first version of the profile was published.

Unfortunately, some people in the community, in the gTLD space, where not happy about it and submitted a request for reconsideration of this profile. So, it got put back in for revision.

Then, ICANN received a proposal from the gTLD registries to instead do a sort of pilot of RDAP, and that was accepted. It started in September 2017. Since then, we have been working with registries and registrars and a few other entities that are interested in this.

I think later comes a resource slide. On that, you will see the link to the pilot page. If you're interested, you can see who is participating and how to participate if you are interested.

Then, in parallel with this work on RDAP comes the fun stuff about GDPR that some of you may know, especially if you are in the European Union. So, with that, in the ICANN space, comes what is called the temporary specification for gTLD registration data. That specification calls for, among other things, for registries and registrars to implement

RDAP following a common profile, set up service-level agreements, and register important requirements.

So, it calls for ICANN to work with the registries and registrars to define these three things, and then go for implementation.

Then – oh. I guess I should have put these slides in a different order. These are the set of pictures that RDAP has. No surprise that they basically correspond to the [inaudible] that WHOIS has. With RDAP, you can have summarized query and response and error messages. You can have secure access to the data over HTTPS. It can be extended easily, and it enables differentiated access so that users get to see some data, and others get to see a different set of data.

It has a bootstrapping mechanism, so if you need to know which server to query, you just query, and an RDAP-compliant client will be able to find a [inaudible] and give you the information.

It has a standardized redaction and reference mechanism. If you need to do so – for example, if a registry wants to redact or reference the registrar – they can do that, and the [difference] can be followed by the client.

It builds on top of HTTP, which many organizations, of course, know how to do. It has internationalization support from the start, like WHOIS has done. It also enables searches, if there is the [appetite] for that.

But, going back to the implementation status in gTLDs, where are we here? As I said before, the temporary specification calls for implementing RDAP following those three elements that are there. For

the first component they profiled – the gTLD RDAP profile – we got a proposal from the Contracted Parties; that is, the gTLDs, registries, and registrars. We work with them, but it's basically their proposal. We, as a matter of fact, published our input to that proposal, and we published all that for public comment, as is the usual process in ICANN. That went from, I want to say, the 31st of August to the 13th of October. So, it just recently ended its public comment period.

We received some feedback, and we're in the process to analyze and work with the Contracted Parties to potentially revise that proposal and finalize it. There is more on the timeline of that in a couple of slides later.

On the other two components that the temp spec calls to have in order to implement RDAP – the SLA and the registry reporting requirements – unfortunately on those we're still negotiating with the Contracting Parties, that's still pending to be resolved. We are hoping to make some progress while in Barcelona. There rare only a couple of issues that need to be closed, so hopefully we can finish that soon. But, we'll see about that.

So, here is the timeline. So, you can see that the first item there is when the public comment ended. That has already, of course, passed. We are still hopeful that we will have a proposal for the SLA and registry reporting requirements, still in October. We will say about that. That will go out for public comment so that people can provide input on that.

We were first thinking that we will need to have the three elements finalized before requiring implementation. Now, we're thinking that

maybe we can request implementation only with the profile, which is the most substantive of the documents in terms of the requirements that had to be implemented.

So, we're thinking that probably we can finish the profile somewhere between November and December. I said that [this conversation has to go with – the Contracted Parties have submitted] a proposal on what changes are to be made.

So, if we go with that, then, per the temporary specification, the registries and registrars will have 135 days to implement, which will take us in somewhere at the end of the first half of 2019. So, somewhere, say, between April and June next year is when we could have the service in the gTLD space start to become available.

In parallel, we expect to finalize the SLAs and reporting requirements. Hopefully, before that, the service becomes a requirement.

So, that's the tentative timeline. I cannot emphasize enough that is a tentative timeline for the implementation on RDAP. I've been presenting these timelines for quite some time, and it's looking every time closer and closer to how fully will get to this one, for sure, next year.

So, in terms of next steps, I already covered what needs to happen in terms of the profile. But, leaving aside those three elements, the next step – what we see now coming in will be the first phase of RDAP. For example, part of the temp spec – it's only covered in the minimal output requirements that need to be covered.

But, doesn't talk, for example, about how to have a consistent or a unified access model for differentiated access. That's something that, in the context of ICANN, is, as you can imagine, a complicated discussion. It's still ongoing and probably will take some time to be resolved in the community. But, that's what we see as the next phase of the work in regard to RDAP.

For what it's worth, with the pilot that we have ongoing with the registries and registrars, we have already talked with them with the intention to continue that pilot to experiment with, in this case, the authentication/authorization technologies that could be used for this unified access model.

We are, at the moment, thinking that it could go in parallel. We will have a technical discussion going in the context of them. The pilot – the authentication/authorization technology could be agreed to with the Contracted Parties and the other interested parties.

In parallel, the rest of the community – perhaps more like the lawyers and the policy people – will discuss what are the requirements in terms of how that unified access model would work. So, who gets access to what portion of data, basically. And, that's put in simplified terms.

So, those are the next steps in terms of differentiated access.

On the client side, this is something that is of a lot of interest for ICANN. With RDAP, as some of you may know, the protocol was assigned to be not exactly human-friendly. The intention of RDAP is to be something

that can be easily parsable, summarized, etc., so it's something that machines can understand easily, but humans? Not necessarily.

So, in that sense, we have thought that there are three different types of users. This is an arbitrary definition. I'm not intending it to be the true one. But, we see that there are the very technical, frequent users. For them, RDAP as it is is enough. If someone is technical, then they can just use RDAP. They can develop their own tools. They can plug it to their own products, and it will work good enough.

Then, the second category we define as the technical users that are not so frequent. So, they will not be [experienced with] how to deal with that. They probably can deal, still, with some command line tools and some, maybe, JSON formatter. Maybe that will be enough. Those tools are already available, so we're not so worried about that.

But, the third category is the one that is of most interest to ICANN and more of what needs to happen in order to make it work. That's the common user, the one that doesn't know much about technology. They just want to get access to a specific piece of information.

So, in this case – this is where ICANN is interested in providing at least something – as some of you may know, we already provide a tool in WHOIS.ICANN.org. People can send queries for gTLDs, and it presents the output in a human-friendly way.

So, we are already working to extend that to provide similar human-friendly output for RDAP. We already have a prototype working, and we

are now working to make it to production. We expect to have that in the next coming months.

One of the things that we have identified is that we're playing with Java script client with the intent to make that client run from your laptop. It will not work from the ICANN servers, for example, so that ICANN will not see what is being query, what is being responded, or the credentials, once there is authenticated access.

In that sense, there is potential for some requirements needing to be put into the servers – for example, that they support a specific [inaudible] recommended but not required in the RDAP standards. There is reference there. The access control allows [inaudible]. That's probably something that we are – well, that's not probably something. It's something we are proposing to the Contracted Parties that should be included as a [inaudible] requirement for the servers, at least in the case of unauthenticated access, so that Java script clients can work and we can more easily provide this to the common user in a human-friendly way.

For authenticated queries, I think it's too early. It depends on what technologies are chosen at the end; what will be needed in order to make this work.

This is the final slide. These are some resources that you may find useful. We have an RDAP page on the ICANN website with all the links to these things and a few other resources for end users, registries, and registrars, and some other stuff.

There is a pilot page, which you don't really need all these links for. Just the first one will give you everything you need. But, I think it's important to mention that, for the pilot page, we have one. There are a few [inaudible] that are playing there with a number of gTLDs.

There is a mailing list that's an open mailing list that's available for anyone to join and talk about not just this topic but any technical topic in the gTLD space.

I believe that's all I had for my presentation.

EBERHARD LISSE: Thank you very much. Any questions?

Thank you very much.

FRANCISCO ARIAS: Thank you.

EBERHARD LISSE: Okay. We are a few minutes ahead but not enough to have a coffee break. So, Brett Carr from Nominet will talk about Anycast in the cloud.

BRETT CARR: If I talk quickly, don't forget a coffee break.

Good afternoon, everybody. So, it seems I stand in front of coffee. This afternoon I wanted to talk to you a little bit about what Nominet has

done in regard to Anycast in the cloud in relation to our DNS infrastructure.

So, I'll do a little quick introduction of me and my team and Nominet, a short history of our DNS infrastructure, and how we've thought and addressed expansion of that. We'll talk about Anycast in the cloud and how it's simple and cost effective. We'll talk about some issues that we had during our deployment and where we're moving forward with this.

So, first a little introduction. I would imagine that a lot of people in the room know who Nominet are, but for those people who don't, we are the registry for .uk domain names. We're the registry operator for a couple of other gTLDs. We're also the backend service provider for 30+ gTLDs

As I said, my name is Brett Carr. A manage a DNS team at Nominet, but I've got a background in DNS engineering – less of a background now [inaudible] because I'm forgetting everything.

Other people involved in it – I'm not going to list their names, but the names are there for you to see, because I want to make sure that they get some credit, because all of those people are involved in this project.

So, a little bit of a potted history of Nominet's DNS infrastructure. I've worked for Nominet on and off for ten years. When I first started about ten years ago, we had seven Unicast nodes. Mainly they were located in the U.K. There was one in Europe. They were all physical infrastructure, a mix of Hewlett Packard and some Solaris servers. They'd done this pretty well, but by the time 2015 came up, when replaced them, they

were kind of creaking at the seams a little bit. They were still performing but they were not long for this world. So, we knew we had to replace them.

This was a time when we were only running a limited amount of gTLDs, but we had signed a plan for expansion into the gTLD market [inaudible], and we wanted to build a more agile, more future-proof infrastructure.

In 2015, that was when we designed and built out a new fully-Anycast infrastructure, which was eight nodes spread across the U.K, a couple in Europe, and a couple in the U.S. So, it's eight physical sites announcing four name servers.

This was based around on-premise virtualization infrastructure, so we had some physical servers in each location, but they're carved up into VMs, certain VMs providing certain TLDs, etc., etc.

We've been pretty happy with that. It has held up very, very well. We've had a couple of large DDoS in that time, and that's not been an issue. It's way, way overprovisioned for what we need, but that's an [itch] of running DNS.

So, earlier this year, we decided we wanted to expand it a little more. We're getting more globally spread customers. So, there's need to have a more widespread infrastructure in more places to try to cut down some latency to some clients. It's partly about cutting down latency and partly about expanding the ability to potentially [swallow] further DDoS attacks and try to localize the impact of those.

So, the obvious thing to do is this point is we've got a [inaudible] built ten [inaudible] twelve, but, as you can imagine, building those physical infrastructures in more and more locations is quite expensive. So, we tried to think out of the box a little bit and think that we could do something different and not do the same as what we were doing in the past.

The cloud is this thing that's in vogue these days. So, we used at looking other people's computers, which is cheaper than using your own computers.

Now, back in 2015 – I guess it hasn't really changed that much since 2015 … well, I guess there's more availability of cloud infrastructure nowadays than there was back in 2015, so it wasn't really a choice in those days. But, when we looked earlier this year, the obvious content is where Amazon, Microsoft, and Google – there are a lot of smaller places as well in the marketplace that will provide – essentially, what we're looking for here is as much infrastructure with a service with some APIs to glue it together, rather than actual, real cloud services as such.

So, we looked to AWS and Microsoft and Google and looked at some of the smaller places as well. But, we already had some skillset in AWS, so we actually selected Amazon as the most suitable way to move forward and started to do more research in that area.

Fairly soon in that process, we found a couple of kind of fairly major issues with using Amazon. Firstly, obviously we want to do Anycast. The support for Amazon to use your own IP space and bring that in and

announce it is non-existent, basically, without some special stuff that we'll talk about in a little while.

One of the things I liked about Amazon when I first started to look at is the auto-scaling functionality, where you can put servers behind an elastic load balancer and have them spin up extra resource when you need it.

Unfortunately, that functionality in AWS does not support UDP, which is pretty essential for DNS, really.

So, when we came across these issues, we decided to search for some help, because we're not experts in this area, and rather than try to become experts or hire experts, we thought we'd try and partner with somebody else.

One of my colleagues found a company in the U.S. called NetActuate, which I think – I want here this morning – gave a presentation here this morning as well. We had a meeting with NetActuate, with a view to using something called Amazon Direct Connect. For those who don't know, Amazon Direct Connect is the ability for you to have a private connection into AWS to either your own infrastructure, or, in this case, your partner's infrastructure. So, what essentially what you can do is run VMs in AWS and then have a private connection coming out of AWS to somewhere, either to your own infrastructure or somebody else's, and then announce your own IP space, hence bringing you Anycast functionality, potentially.

But, when we got talking to NetActuate, it was apparent quite quickly that they had a lot of experience in Anycast. So, we were quite impressed with how authoritatively they talked about this technology because Anycast is [not] quite well-known in the DNS industry. And, it's becoming more well-known over the years in other parts of the networking industry, but in a lot of places, it's not very well-known. So, we were quite surprised about how much knowledge they had.

On further chats with them, it appeared that they had infrastructure in 25+ locations globally, which is actually more locations than AWS. That infrastructure is a service location, so they can provide virtual machines in those locations.

They also provide some quite good, solid experience with other customers running DNS. So, they already had people doing DNS-Anycast-related stuff in their network.

Their infrastructure had very good API access. One of the things we wanted to do is try and alternate this stuff as much as possible, so that was an attractive part – what they offered. Their pricing is as good as AWS, and, within their network, they had some built-in DDoS protection.

So, we decided to move forward. So, instead of using the AWS via NetActuate's method to actually deploy virtual machines within NetActuate, we selected four locations, them being Dallas, to fill in some space in the U.S. that we didn't really cover very well. Of our physical nodes, we've got one on each cost of the U.S. but nothing in the middle. So, we put on in Dallas. We put one in Sao Paolo because

we have nothing in South America. We put one in Hong Kong because we've got nothing in Asia, and we put one in Sydney because we've got nothing in Australia.

Those locations were based on some research that we did on where some of queries came from. I have some analysts on my team who would analyze data, and that's where we suggested we put our first nodes. But, the beauty of being able to do this in the cloud and with somebody like NetActuate is that we can do it. There's no [inaudible] at all. It's all [inaudible], so we can do it. If it's no good, we just stop it and we move it somewhere else. So, it provides really good flexibility.

We put one virtual machine in each of these locations that serves all of our zones. That's different to what we're doing in our own locations, where we have separate VMs for groups of TLDs. Each of these VMs is fairly chunky, at eight virtual CPUs and 32 gigabytes of memory.

We've run some software, which is a lightweight BGP [inaudible] called ExaBGP. We've basically [set] a peering session directly with NetActuate [inaudible] and it announces our prefixes to them.

One of the engineers on my team also wrote a health checker script, which basically runs constantly on the VM and checks that the DNS zones are being served correctly. If it detects any TLDs not being correctly or not being served at all, etc., it will automatically withdraw the relevant prefixes within ExaBGP.

One of the things we were quite conscious of doing that we wanted to do quite well in these sites was to be able to offer some kind of filtering

capability so, if we get DDoS with actual DNS traffic, where we can identify patterns in that DNS traffic, we've got somewhere to filter out what we would deem to be bad request. So, we decided to do dnsdist for that. We do a different thing in our nodes for filtering, but, because of where that's implemented in our nodes, that wasn't available for the cloud nodes. We used dnsdist here as well.

We use nsd as a name server in these nodes. Some of you will know that we have a custom DNS statistics tool called Turing. We also have the Collector software collect queries at these nodes and then send it back to our data warehouse.

One of the things we wanted to do, like I said earlier, was make this infrastructure as automated as possible and to involve as little input from engineering as possible. So, we don't want to build lots of lots of cloud nodes and then spend lots and lots of time patching them and messing around with them and looking after them, like you've got to look after kittens, for instance, as they grow up into cast. I should have a picture of one of my cats here, I suppose. But, never mind.

So, what we decided to do was start to use something called immutable infrastructure here. Basically, that means we operate with a single image for all of our nodes. Whenever we want to do an update to any of the infrastructure, we update that single image, and then we kill the rest of the infrastructure and rebuild it [and] the entire infrastructure from scratch.

We use something called Ansible, which is kind of an automation and orchestration tool set, and basically, some Ansible scripts that will

rebuild all of those four nodes from scratch. Hopefully, it'll scale to lots more nodes.

So, we don't spend any time patching all those servers. We just maintain that single image. The VMs are born, they're used, they're killed, and they're reborn again.

These are just a couple of slides to illustrate what I mentioned earlier, that NetActuate has got some DDoS capability. This slide just shows novel operations. So, the green there is client DNS traffic coming into the NetActuate data centers. They have a DDoS platform that samples the traffic and decides what's normal and the traffic then appears on the four nodes.

If they detect DDoS traffic, they will redirect that traffic into one of their – I can't remember – I think it's four global scrubbing centers and scrub the nasty stuff out and then re-layer the good stuff.

So, we did run into a few problems during this implementation, which I thought I'd go through quickly in case anybody else does something similar. When we initially got things up and running, we had less traffic than we expected, so we had to work with NetActuate to kind of do some routing tweaks and changes to bring the traffic up a little bit.

When we did that initially, we actually ended up with more traffic that we wanted, so we ended up with the cloud nodes doing as much traffic as our large global nodes. Now, it's not what we were trying to aim for. We wanted to make a good balance, with the bigger size getting more

traffic and the smaller size being more local traffic going in through local peering. So, we did some [inaudible] to bring that down again.

During our testing, our global sites can do more than 500,000 queries per second, and the cloud sites can do about 100,000 queries per second. So, we need to be aware of that and take care, and if we do see a large DDoS, balance the traffic correctly.

Lastly, one of the issues we had for a little while. We would, in normal operations, when everything was fine – we've got some monitoring nodes that are spread all over the world. We have both UDP and CCP tests going on in those nodes. Periodically, we saw some spikes of TCP latency for some unknown reason, so it quite a while for us to work out what was going on here. We did a lot of work with NetActuate, and they were really, really helpful with this. I think there's two of them in the room, so thank you very much for the help, you guys.

It turns out that this is an issue with TCP offload, which is common in the Xen and KVM hypervisors. To solve that, you need to turn off offload on the VM itself. We'd never come across this in our infrastructure, because our infrastructure runs VMWare, and VMWare doesn't suffer from the same problem.

So, for the future, we are planning to roll this out into further sites at some point. Probably I'll run it next year, I think now. But, we are currently also building some recursive infrastructure in the same platform, in NetActuate, using the same methodology and the same image, etc.

That's it. Any questions? I'd be glad to answer.

EBERHARD LISSE: Thank you very much. If there are questions, ask them now. If there are not, I'm willing to give to you a ten-minute coffee break.

BRETT CARR: You can ask me questions in the coffee break as well, if you want.

EBERHARD LISSE: Okay. But, the next presentation will start at quarter past. Okay?

BRETT CARR: Thank you very much.

[END OF TRANSCRIPTION]