DUBLIN – How It Works: Domain Name Registry Protocols
Monday, October 19, 2015 – 14:00 to 15:30 IST
ICANN54 | Dublin, Ireland

ED: But I'm with the US government. Well Steve here is, okay, doing his job. Okay. So, this talk is about the protocols that are used in running a TLD registry. And from the questions I've gotten so far on this, these protocols, I'm trying to key on the protocols. I want to talk about how things are sent back and forth.

A protocol is a way of talking between two places. This is also common to registrars, but all of my materials pretty much relate to registries. You can ask a question about registrar activities. I will know how they work, but I don't have as much background in some of the day to day businesses issues that are in registrars.

So let's start with a registry. A registry, if you go back to the dictionary, a registry is an official list of things. It generally will say this thing is going to be associated with that person. That's a general way that registries are used. Now in this tutorial, to go quickly from a very general registry term, we're going to go to domain name registry, where we have domain names, are things that are registered with people in mind. Next slide.

Note: The following is the output resulting from transcribing an audio file into a word/text document. Although the transcription is largely accurate, in some cases may be incomplete or inaccurate due to inaudible passages and grammatical corrections. It is posted as an aid to the original audio file, but should not be treated as an authoritative record.

Now within domain name registries, you can have domain name registries in many different places inside the DNS, specifically and I'll talk about top level domain name registries which generally are the registries that you hear the most of within the ICANN context. So what we're saying here is not specific necessarily to what happens at ICANN, but I'm going to try and tie it back to the ICANN terminology, ICANN activities this week. So to first of all show what other kinds of registries are out there, to give you a mind about what registries are, RIRs are registries.

They don't do domain names. They do network addresses, and network routing information, and AS numbers. There are protocol parameter registries in the IANA has a whole lot of those because what does a certain number mean for a certain field and certain protocol? And outside of the Internet, there are things like land ownership. If you own land somewhere, it's registered.

Someone keeps track of who has what plot of land. If you have a car, someone has to keep track of who has that car, license plates will help people identify the driver of a car, or at least who owns the car. And finally, gift registries are out there too, that's a much more different way of thinking of registries, but someone says I am having a baby and here are the things I would like to have, and people say, well I'll give you this, I'll give you that.

Again, what gift is kind of associated with a person so there is no overlap on that. Now in the DNS tree, we have a couple of different registries of note, at top, we have a root registry which is the one that's operated within IANA, which is the starting point for just about every DNS action taken.

Below that, there are a couple of registries, they all act as TLDs. They are subdivided for different letters, gTLD is the generic TLD. Another version of that is an STLD, a sponsored TLD which has become less common over the year. But gTLDs are basically business TLDs. They're run under contract with ICANN, for the most part. Another kind is ccTLD, country code TLDs. Again, these are TLDs which in the protocol are the same, do the same thing. But their administration is totally at the whim of some other body.

Usually the so-called country code they're associated with, whether or not it's actually a country. IDN ccTLDs are also top level domains that are run under the country code operations structure, or manner. But in this case, the IDN infers that the name is not just two letters, or two ASCII characters, it's whatever script is being used for that particular country or region.

And finally, other TLDs. This division isn't quite, it doesn't capture everything. There are a few other odd ball ones in there.

RPA is a TLD, but that's just a whole address of parameters. It's not really a registry in the sense that you register names in there, but it is a TLD also.

And below each of these, generally are private organizations that can do things. Some TLDs are open for registration, and anyone can put a name in there. Some have restrictive rules and some have rules that says no one gets to put a name in the area. The name is being held just for some specific purpose.

So service definition, a service definition of the TLDs. In order to talk about the protocols that a TLD uses, and hence what is speaking, you want to know who it's talking to. And that's what I want to spend a little time on this in the next couple of charts. TLDs will have basically at least three faces to them. The first face is going to be those people that are going to register objects in the TLD.

If you've got a domain name, this is my domain name. I'm going to go pay money, I want to be in the system. You are called a registrant. In the ICANN world, you are a registrant of a domain name. Registrants talk to either resellers up in the top, the little box up top, or strictly as a registrars to the bottom part of it shows.

And there is an entire body of works just on the registry between a registrant, the registrar is out there, the resellers out there, but then put the names into the TLD registry. On another side of this, another face of this area is the support given to TLDs to make them a little bit more trustworthy, a little bit more compatible with the environment that they're in.

In this case, I've thrown in something called data escrow area. This is a place where, that the data that's being held, the register that's being held inside the registry, is put in safekeeping for emergency purposes. And finally, and also down there is the Trademark Clearinghouse, which as trademark holders have a say in whether or not names or registered in a way that might be abusive to their concerns.

And next. And then the third side of this is the general public. People who don't have anything like a domain name, I've actually never purchased a domain name in my life, I've never been a registrant. I've never actually bringing in the DNS, but I go on the web and I use web browsers, and I send email back and forth, I'm going to be using domain names in what I do.

I type someone's name, someone's name at a domain name. People like me see the registries through the fact that the registries are going to provide me DNS information. So that's a whole other crowd. Now this is the crowd that sees the biggest impact of a TLD going bad. If a TLD goes down, this side of their business goes down, disruption happens throughout the

Internet. This is the side that also contributes no money to this problem here. We're just looking at the information.

It's kind of [inaudible] thing. The registrants are funding these TLDs. They don't see a whole lot of interaction in terms of what has to happen day to day. Those of us who are free riders are making them do all of the work. Next.

So in this talk, I'm going to talk about these protocols that are arrayed around here, and there is no particular order to the picture in the following grouping. EPP is one protocol that you may or may not have heard of. I'll explain that. I'll explain all of these as we get to them, but I'll just say the names right now. DNS and DNSSEC are two protocols that are mentioned all of the time here.

There is a WHOIS protocol, which is also gets a lot of talk. The R-DAP, which is kind of a newer thing. Data escrow which doesn't really have a name, I just kind of call it what it does. And then the trademark area is another area to talk about. So let's start with the DNS, the Domain Name System.

The DNS is pretty much the premier protocol on all of this. It's a look up protocol. It lets you say, "I want to know the address of this host." And the host is a domain name. I look it up. I get back the address and I can go there. It's a query for information protocol. The response will come back, it will either be the answer, ultimately be the answer, it just will be no, there is no such host or there is no such address, or there is no place to go. Next.

It's a fairly significant protocol. It's one of the oldest ones in use. It's not the oldest, but it's one of the older ones. If you spend time, like I have, building protocols and striking them, and trying to go back and research them, it's really, really, really old. This impacts everything about the protocol. It wasn't to expand that much. It was built to be very fixed sized, because the time this was built, fixed sized protocols was all we can actually handle.

We didn't have a way of handling modern, fancy protocols. This has made it very hard to replace the DNS, so we end up suffering with it. But if you look through the history of technology, we have so many cases where something was done once, not done correctly, but we've learned to live with it to be just as it is, and that's where the DNS is right now.

It is the reason for the registries to exist. Now next.

Now, it's the, in terms of resiliency of TLD, that's always available, the DNS is the one protocol we talk about. If we look at any of the SLA agreements out there, the service level agreements that are out there, the DNS has to be 100%. There is no excuse for the DNS to ever be out for a TLD. We should be able to have it up all of the time.

Other components, other things can go down for a while. The registration system may go down for a while, some of the other backups may go down for a while, but when it comes to answering people's questions, there is no reason why you can't always have an answer somewhere.

It also is the one component where we have no idea how many people are using it. That's the most frustrating thing with dealing with DNS is that, I know who is running DNS, like I can go through all of the TLDs, I can enumerate them, there is only like about 1,000 or 1,100 right now, somewhere in that area, but I don't know how many people are using it because there is just no way. There is just so many ways things come into us.

Next. In a block diagram form, the orange bubbles show the registrar, which is the external element here, and also IANA. They impact the DNS by the registrar constantly hitting the registration interface, which I'll talk about later, through the registry database, to the server which speaks the DNS protocol. IANA in this picture is there because we configure the TLD at one time with them, and we just run with it from then on.

Next. Components. The DNS is a, it has a few pieces that I want to talk about. It has an authoritative server. This is a server of DNS information which says, you ask me a question, I'm going to give you back an answer that I was told by the administrator, there is no way this came from anywhere else. It came from the real source, the one true source for this.

The second source of answers, when you ask a question, something called recursive servers, or caching servers, intermediary servers. They're middle boxes basically. These are servers out there that learn from authoritative servers information repeated back to someone when they ask for it.

It's not the authority, but it's a copy of that somewhere. And finally we have stubs and clients. These are the sides that ask for information. Every computer out there that wants to lookup a domain name, has a stub of the DNS in it. And it will ask some server, probably a recursive server, that will then ask an authoritative server if it doesn't know the information. Next slide.

And so another way of looking at the DNS and how it effects all of this, between the DNS server, the blue box on top, it speaks DNS to the recursive servers, and the DNS again down to the clients out there, the machines that you have in your hand. Next slide.

Now DNSSEC is the next part of this. Now DNSSEC isn't a separate protocol. DNSSEC is a very significant addition to the DNS protocol. But it gets, it always gives its own set of slides.

It's that much work. Because of that intermediate step in there, the DNS system is very gullible.

So we had to come up with a way to secure it in some way, and DNSSEC came out like this. Now, one of the first ways to secure something is to say, well I'll have end to end security. I'll make sure that whoever I talk to, I talk to through a secure channel, and I know who I'm talking to and back and forth, but that doesn't work for DNS because of the intermediary points.

We couldn't do it that way, we had to do it something else. Next slide. The history of DNSSEC. It was first proposed as early as 1990. There was a first time that we noticed that there was a gullible part of the DNS. Until 2004, it went through a series of protocol developments and a lot of workshopping for about five years, the last five years of that time period. One of the reasons it has been working its way through the system is that it was designed with the operators trying to bend it at the last minute before it formalized everything.

In 2004 the IETF issued a set of documents, defining the protocol. It has been updated since then. And then for a couple of years, it just kind of sat there, because it's not easy to do this. And there is no real need to do this. It doesn't do something exciting. I've always likened DNSSEC to being seat belts in a car.

If you get in an accident, they could save your life, but until then, who cares about them?

You know, you're never going to buy a car because it has good seatbelts. No one designs seatbelts. No one decorates their seatbelts. It's not very exciting. But this guy, Dan [inaudible], he's a researcher. He was not an attacker, he was a researcher, demonstrated how vulnerable the DNS was in 2008, and showed how to attack the DNS from a theoretical point of view.

And that spurred a lot of interest in DNSSEC, saying we've got to stop these kind of attacks. What he described that's happened is called cache poisoning. I could forge answers and shove them into the DNS protocol stream. And suddenly I could make a resolver give out bad answers for something for quite a long time.

So DNSSEC took up some steam at that point. In 2009, it first appeared in Sweden, the dot SE zone, with the first to rollout a signed zone. By middle of 2010, the root zone was also signed. It has been signed ever since then. Now the approach that's used is digital signature. I take some data, I take the data that DNS would answer back, and I apply a digital signature to that.

I used public key cryptography, which means I send out public keys, you all get a copy of the public key. When I sign something with a private key, that only I have. It's a secret known to me, if you could verify that that signature over the data, you know it had to come from me.

That's the idea of public key cryptography. To make this scale to the global size it has, we use a hierarchy of keys. The key for a piece of data belongs local to that data, is chained all the way back up to the top of this. There are people who don't like the idea that DNSSEC is a centralized security system, but frankly it's protecting a centralized naming system, the DNS. So it follows the DNS tree.

Now what a registry has to do, first of all it has to manage its keys. If you are doing any signing of anything, whether it's PGP, encrypt your email, DNSSEC, or SSL certificates, you have to start managing your cryptography somehow. That's a job that is often overlooked, a job that gets fallen off sometimes, if one loses interest in it. But you now are, your job is to manage these keys.

The next thing TLDs do is that they have to now accept data from their registrants, that's DNSSEC, applies to DNSSEC. In this case, they're called the DS record which is just a particular record that DNSSEC is setup the chain. The TLDs have to sign these records, to sign the DS records and publish that, and the next bullet I have up there, the next to last bullet up there, signing negative answers.

That sounds kind of simple, but in the DNS it's actually kind of hard because in typical old time DNS, when there is no answer for what you want, I would send back an empty message. You ask for something, you get back nothing. You can't sign nothing. So you have to sign something, so we had to create something to sign, and that something has been contentious over the years.

And finally, as the TLDs, when they start managing their keys, have to make sure the IANA, one step up, knows what keys they're using, or else the chain of trust will get broken. Next.

In bold diagram again, the registrant sends in information that the DNSSEC functions part of a TLD, and a root registry is also interacting with that function there. It goes through a database, populate the database, which is the registry of all domain names and their information, down to the server and out is DNSSEC.

Next slide. And this week, there are two sessions dedicated DNSSEC, which you might want to look at. Later today at 5:30, about three hours from now, there is a DNSSEC for everybody, a beginner's guide, that's going to happen in, right here. So don't go anywhere.

And then Wednesday, there is a much longer session, about six hours, more talks about DNS workshop. Again, it's right here, so don't go anywhere the next two days. You wasted your hotel room for the next two nights, you're going to be here until then, right? Next.

WHOIS. Oh, the protocol. WHOIS is a protocol. It's older than DNS. It was necessary at the start of the Internet because when you put a bunch of unreliable pieces of software and hardware together, which is what this wall was in the 70s and 80s, you had to be able to call up somebody on the telephone system, which was then was actually working really well. And say, "I can't get to you, what's going on?"

Because you don't know if a problem is going to be at your end, at their end, or somewhere in between, and you have to work out this way. So in a sense in a day when everything was funded by a certain set of agencies, research across a couple of universities, there was no money riding on this thing, no enterprises, no investment, just a bunch of researchers out there, you had to know everybody for this to work because we're building in a virtual world.

The WHOIS protocol, a very simple question and answer, again, there was no concerns at all about security, privacy, whatever at that time. Next. The protocol definition is pretty much what's on this slide here. You open a TCP connection to a certain port number, 43 is the number for that. You send a question. You wait. You get back an answer. You then close the connection and you print the answer to the screen.

There was nothing at all said about what the question was. Nothing said about what the answer is going to be. Next. And the way that fits here, the registry database feeds the WHOIS server, the WHOIS server then speaks WHOIS out to the clients.

Now, why is there a problem with WHOIS? Well, questions and answers are undefined. When you're trying to write software, you can't really design to whatever. It's very hard to write whatever, it's very hard to parse whatever. Freeform is not good for interoperability out there. If you ask one registry for the information about a domain name and say, who is the owner? It might say, well no one owns them here. The person who is currently holding it is this, as opposed to the organization, the administrative, all the different terms that are being used for someone having the name could be used.

Another biggest problem back then was that everything was an ASCII. It was assumed that only ASCII was involved here. So there was no internationalization at the time. Now for someone who is trying to manage a system, this first protocol assumed that all of the information was in one server out there, something knew everything about everybody.

We had no way to redirect you saying that, I only know about this part of the name space. Go over there and ask somebody else. The DNS does that natively, but it wasn't in WHOIS. The idea of differentiated access, that's the idea that law enforcement may want to have more access than a general person out there, or that I may not want to give any access out to this for certain kinds of…

None of that exists in the WHOIS protocol. There was nothing at the client end that could support any of that. And there was no means at all to validate the data inside these answers. Whatever was registered there was registered as is. In fact, over the years, as people trying to clean up the WHOIS area, they find even basic postal address structure was not followed in many places, where you may have someone listing their town, but not their province and not their country or not their postal code.

And in some cases, I've actually gone through some WHOIS data in the past, and I found two domain names. The only thing common between the two of them was the mobile phone number that tie into the same person. Same building, same place, but they put it in different ways. You couldn't tell the mobile phone numbers that was common between the two. Next slide.

And WHOIS, despite being the oldest thing out there, it has two sessions. You notice DNS had no sessions. We're done talking about that apparently. WHOIS review team international registration data expert working group, that meets… Time travel is required for this. They met yesterday. So forget about that one.

When I gave the talk last time, it was at the same time as my talk, now it's in the past. And then also the thick WHOIS policy implementation meeting is Wednesday morning, 8 to 9, and that's somewhere else. So you don't have to stay here for that one. Next.

Okay, so EPP protocol. Now first of all, I'll say the DNS and WHOIS have been around forever in the registration game. DNSSEC predates EPP kind of, but EPP came along about 15, 16 years ago. It's what we call a business to business protocol, meaning that this is a protocol that not everyone does. DNS, WHOIS, everyone out there does those protocols.

WHOIS if you care about something, DNS because you're using any other application out there. But EPP is usually done only between two people who are doing business about the registration of domain names. Its purpose is to edit the register. Update names, update information about names, transfer names between two places where it gets kind of complicated, and any other maintenance that could be account information, plus it's extensible, it does everything else that you want to do within the registration game involving business policies, retrieving stuff and so on. Next.

The history. Between 2000 and 2003, it was developed in the IETF. It came, it was spurred on from some of the work being done in VeriSign in common net, back in those days. They had a previous protocol called RRP. This was an improvement on that EPP done, and encouraged to support the shared [inaudible] model that ICANN has been following for so long.

Between 2003 and 2009, within the IETF, it went from being a proposed standard, or a basic standard, the earliest formal standard, to be a full complete standard in their naming system. It was, it has been mandated for gTLDs and sponsored TLDs by ICANN. It's in all the contracts, I believe all of the TLDs that ICANN oversees does it, but there may be one or two that haven't gotten in there yet.

I just don't know if, about a few of them. Now as far as ccTLDs, which are not, they don't answer to ICANN for anything like this, they have been adopting it. Many have seen it as a good idea, and have brought it in house, and they have been developing on that for quite a long time now, so at this point, EPP truly is a standard protocol from the IETF doing a very generic function, whether it's the ICANN world or outside of the ICANN world.

Currently the IETF is a working group that's been spun up again because in the 2000s, there was a bit of a spread in the way it was extended, and people lost track of how to do it correctly. The IETF thought it was time to bring it back together and thus reduce the number of extensions, and try to keep things more common. It doesn't stop people having different attentions for the same idea. They still exist, but now there is a working group they're trying to encourage a more common way of going forward, whether it's not the ICANN model, the ICANN regulated model of TLDs, or it's the ccTLD doing what it wants to do.

EPP doesn't have to be an exclusive way of updating a registrar. Many TLDs will have other ways to put things in there, if they're allowed to. Sometimes you're required to do EPP, but you don't have to do EPP. It's just one way to change what's in the database.

And the architecture here, EPP uses a very strong transport layer, a secure transport layer. It's done in XML, the extensive markup language, which is kind of an older way of writing these protocols. You have to write a client, along with the server out there, it's… The way things were done, when this was minted, this was protocol was made, this was the way people like to do protocols.

Servers inside the registry, it's that middle blue bubble there, and it speaks it to the registrar out there on the outside. That's how we update the databases. Next.

R-DAP. By the way, if you have any questions, go ahead whenever you want to. Steve will watch for hands, because I can't see that far. Put your glasses on. Anyway, at the end we can also do questions too, because frankly, I call this a laundry list presentation. I'm going through one paragraph to the other one, it kept, after about four or five, they're really boring to do this, because they're not related to each other. So it's not like a long story.

So if you feel want to break it up with a question about something, go ahead. R-DAP. R-DAP is a registration data access protocol. No domain name registry that I know of, though there may be some, uses this right now in operations. It is, there are places operating this in the RIRs, it's a new thing. It's a query response protocol to look at what's in the registration database.

And if you're paying attention, that's the same thing that WHOIS is. Basically, when you're registering stuff in the database, you want to ask what's in there. The DNS is a report from it, WHOIS and R-DAP let's you ask, what do you know about something, because sometimes names are registered, but not in the DNS.

I can register a name for the purpose of blocking it from being in the DNS. I may have more registrations than actually delegation of the DNS. Now, R-DAP had a couple of goals to meet in its design in the last couple of, three or four years in the IETF. I want to be able to ask any server and essentially be led to where the answers is going to be.

I want to be able to go from server to server. And it's also very interesting that this protocol was developed not for domain names. Domain names were not the first thing to be put into R-DAP, the RIRs started this. Numbers were the first thing, numbers, IP addresses, and AS numbers were the first thing to go in.

All this protocol sits on top of https, secure web transport basically. WHOIS, so connect to port 43. You may hear port 43 WHOIS, that's native WHOIS for those who did it for a long time. There is port 80 WHOIS, which is you ask the WHOIS question over a web interface, but those two, that's just a frontend for WHOIS client.

R-DAP doesn't have a port number. It rides totally on top of the web. It's part, it's web technology you go, so web server. It just gets handled internally as a URL. Next slide.

UNKNOWN SPEAKER: But just to clarify that. The https has a port number, so it's inheriting the port that this other protocol is using, right?

ED: Yeah. And actually, that number is 443. That 443 and the 43 are not at all, it's coincidental that they were there. So. Somebody might want to read into that, but they have nothing to do with each other. [Inaudible] R-DAP is a server, it's software to parse the queries. It looks it up in the database, and it prepares a response. It's all that does, pretty much. A client is a web API. It's a web browser. It's basically just following what the web browser already do.

The important thing here is that clients in R-DAP as opposed to WHOIS, the clients are smarter. They can do things. They can support authentication. They can say, I have a certificate proving I am who I am, and I'll present it to you for whatever purpose.

Now the history of R-DAP, where it started from, there was dissatisfaction with the WHOIS, which we all know about. We've probably heard about that. Two of the RIRs out there said let's do a web based approach. They actually tried a few approaches that didn't work so well, but they finally get one that was really successful.

So whenever you get into R-DAP, if you start looking at this protocol any deeper, you're going to see a lot about the RIRs, the numbers side of this, the IP addresses, and also you're going to see a lot about the http protocol and https protocol. And you, a lot of it is just taking over by that in the lower layer. Next.

The query looks like a URL. These have a way of structuring the query in there. It's a formalized way of giving the query, it's not just a generic, it's not a generic question. It's a very structured question. Now the key though is, what I asked for can change over time, but I have to define what it is. Like I said, this is an address, this is a name, this is a contact. I can change it, but I have a fixed number of things in there I look for.

It makes it a lot easier to parse this. The response comes back over to GPS, meaning that all security is taken care of in places like that, they have a bunch of different services in there, retries, finding answers. They use JSON, Java Script Object Notation. That's now replaced XML as the way to write things down in the Internet. Again, it's like a WHOIS in some ways, but it's a formalized, it's parse-able. You can machine parse this.

In the old days, when I looked up an IP address, I wanted to know who had that address, if I went to one of the RIRs, I looked up the administrator, I looked up another, it was the owner. Another one was the organization. I looked for specifically terms that were applied to each of the individual RIRs. With R-DAP, they've said, okay, the person who holds this has this name.

This is the label I used for the owner. It makes it much easier to, for a machine to automatically pick this stuff up. And finally, a formative redirection message again, to use the RIR's example WHOIS, if I didn't know which of the RIRs issued the address, I can try any one of them and they'll say, no it's not mine. It belongs somewhere else. And some of the RIRs said, it goes somewhere else.

Others say, you go to ARIN for this address. They knew which one to go to. So again, you know who you are asking, what you're going to get back. So we have this definition out there. Now, what R-DAP provides, you don't have to use, but that gives you a choice. WHOIS didn't let you have a choice.

Now at the bottom of it, to-do, an operational profile. The protocol is defined a certain way, but it doesn't say you have to run it exactly that way. Right now in ICANN, we're working on the operational profile for those registries that ICANN will oversee. And I know of other ccTLDs that are doing their own operational profiles for their specific concerns.

So, I'll go through this again briefly. I think I've touched on this already. Define data model, it makes it a lot easier to handle. Expansion beyond ASCII, internationalization has been a very big part of this. It uses more the Unicode below that. Distribution of data meaning that the data doesn't have to be in one place. I don't necessarily mean it's in one place, which has a lot of impact on where data is retained and held.

Differentiate access. That refers to whether or not I want to have some data available only to law enforcement. Do I want to have some data that's given out to the general public? What do I want to send to different kinds of people out there? Now, of course, you know, it sounds like a simple little problem, but it gets complicated quickly because first of all, define law enforcement.

Right? Each registry that is going to define who they consider to be the ones that they give everything to, who are the ones that they will give almost nothing to, who will they…? And so on down the line. This has to be worked on. This is part of an operational profile too. What's going to be done about how people get access to this data out there. It's wide open. There are many different interpretations to that.

And finally, this is more compatible with current software practices. Next. And the way it looks here, it looks a lot like

WHOIS. The R-Dap client speaks R-Dap to the R-Dap message handler, which is getting information from the database, and the arrow goes one way. Next slide.

And this week, there is one session on R-Dap. It's Wednesday morning, downstairs, 12:30 to 1:45, where they'll be talking about, I presume, the operational profile is the main topic there, although there are other topics in there. I haven't seen the agenda. Steve.

STEVE CONTE: I have a question from online, if that's all right.

ED: Sure.

STEVE CONTE: This is from Patrick Miles, he says, "Ed, you mentioned R-Dap not being used. Why not? Do you see increased usage coming?"

ED: Yes. Yes. Right now, the reason why not, for the most part is people are rolling it out. It was, the RFC is defining it, we're past the approvals of the IETF just in December of last year, so it's like 10 months later. People are deciding that now that I know what it's going to be, I'm going to see well, how does it fit with current data model?

I know one ccTLD that has been doing this, and right now it's just a matter of software development cycles. It's a matter of making sure they define how they want to roll it out. It's just a matter of a lag if it's going out there now. Within the ICANN world, I know that there is a requirement in the registry agreements that say that, when a new replacement for WHOIS is ready, TLDs have X number of days to implement it.

That clock hasn't started ticking yet because not X number of days for the RFCs coming out of the IETF, but after ICANN says, we want to roll this out through all of the system. And what is happening now is making sure that we put the correct rules in there.

So it's coming, it's just a new thing. I'm sorry Steve.

STEVE CONTE: I just want to make sure, okay. We're okay with that. Okay.

UNKNOWN SPEAKER: I'll get over it. Okay.

ED: All right so, moving into the realm of the protocols which are less fancy. Data escrow. The purpose of data escrow is to take a copy of the registration database, which is basically TLDs source of authority, that's the, I'm sorry. The record of authority. It's the thing that they're building their business on is that contact list. And you have to put it with a third party, whose job is to take this stuff and store it so no one else sees it unless some conditions occur.

Why do this? If the TLD goes under, it fails. Someone steps on the database. The whole thing goes gone. Who is going to protect the investments of all of the people who have paid money for names that they now have done branding on top of? Where are we going to have the information?

So we have to have this somewhere else. This mix of registries have… This makes people… Registering a name has some confidence that even if their single point of failure here is the TLD, I have another place where my information is being held. Hopefully, from a technologist point of view, some other operator could say, okay, just give me what's in that escrow, and I can just populate my database and start running right away.

That's the hope. It's quite far-fetched, but that's what we would like to be able to do is to get things up and running again because again, the DNS side of this is going to affect everybody out there, not just the registrants and other parts of the system. Now stored by a third party, that's pretty simplistic, is stored with strict rules of access.

And in this case, ICANN tells all of its TLDs that are under the agreements that say this, you will have an agent that will collect this information. ICANN can get that information if certain things happen, but we also need to have a key to get the information, because it's not going to be in plain text, and I'll get to that soon.

So next slide. The history of this. The IETF had a birds of feather on this, about five years ago when I was a co-chair of that, so I have a real good memory of this. The IETF didn't like it. They said it was boring. They weren't excited. It wasn't a proposal, it was, do you want to talk about it? No, that's not interesting.

That doesn't mean it's unimportant, it just means that it's not that technically interesting to a bunch of engineers that like to do really complex things. It's actually really a pretty simple thing. I just want to take this data, put it in encoding, ship it across to someone who will then put it on their database. I want to make sure it gets encrypted, that's probably the most complicated part.

Well actually it's not the most complicated part. The most complicated part is, I want to compress it. I want to be sure because you've got these third parties charge by the byte basically. So I want to make sure that everything is small when it gets there. And that's actually the hardest part of all of this.

So that's kind of where it comes from. It's defined currently in two, go ahead. In this case, my slides actually matched what I wanted to say and you, I'm sorry. We'll talk later. So it's the fun in two places. If you look for data escrow in the ICANN world there are two places that's defined. Not defined twice. There are two pieces to the definition.

There is what looks like an Internet draft that exists somewhere in the ICANN world, which describes the framework. It describes how the escrow is going to work. What do you deposit, and so on. Secondly, there is the timing of this. When do you deposit this? And how do you deposit this? And so on. It's in specification two of the registry agreements, at least for the 2012 round of gTLDs. It says here that the dump of the database is going to be XML version in one or more files. It doesn't necessarily have to be XML for all of the TLDs. I know that some of them use a different form.

It's encrypted so that it's cheap and it's only seen by the appropriate parties. Every day deposits made. Deposits being the incremental stuff that has happened in the last… What names they sell? What was unregistered that day? And on

Sundays, we have a full dump of everything that goes in there, incremental the rest of the week. And so this chart here, which unfortunately we didn't do our color scheme right, but the registry operator on the left will send to the escrow agent the public keys.

The middle arrow says that the deposit was a full or incremental. And it also lets the escrow agent know that this has happened for whatever purpose of telling the escrow agent. ICANN gets notifications from the registry operator says they've done it. And the public keys are also sent along. The public keys give ICANN the ability to see what's in the escrow agent's, what's being held by the escrow agent.

However, just because they had the ability to see it doesn't mean they get it. That's the difference here. ICANN doesn't have the deposits. They've been locked away in some other place, and will only get it if certain other conditions are met, and then ICANN will talk directly to the escrow agent, which I don't show you because I'm talking about the registries.

Trademark Clearinghouse. This is the last of the laundry list. It's an openly defined way to protect… Yes, a question.

This is where Steve gets exercise. I like this part.

STEVE CONTE: I didn't see where the question came from. Here we go.

UNKNOWN SPEAKER: Hello. I am [inaudible], I am from Jordan. I wanted to ask you about the escrow agent. Is it a sort of backup?

ED: Essentially yes. It's essentially a backup. But they're not able to process the data. It's sort of like backing up your hard drive, but…

UNKNOWN SPEAKER: For how long we keep that data that we encrypted and compressed?

ED: Off hand, I don't know. I would assume in this case, they only retain it until the next full. There is no reason, I don't… I'm sure they're not contracted for the history. You have an idea?

UNKNOWN SPEAKER: If you corrupted the data and you waited two weeks, then there would be bad data and nothing to backup, so I presume it's a couple of years.

ED: Good point. I haven't looked at the detail. So, I'm sure there is retention of some time, I just, I have never looked at that level. But and I am aware that are times where some people will report that they've submitted data, but they sent an empty file too, so there is probably some history kept.

UNKNOWN SPEAKER: Just to bring that up to the point though, that's a contractual agreement issue. There is nothing to technically say, you know, you can only hold that data for 86,400 seconds. That's a contractual deal between whatever parties are contracting, and we specifically are focusing on the technology underneath that. So actually don't want to touch the policy, we're trying to be neutral in what the technology is providing.

Was there another question too? Okay.

Yeah, it's actually, there is three layers. There is technology that there is the contracts, and then there is the policy in this. The contracts I haven't scrutinized for that level of detail, to be honest before. Because other questions came up like, well does the data escrow tell ICANN something like, well I didn't look that up. I'm onto the registries in this set of slides, so.

So, TMCH. Trademark holders had concerns about the opening up of domain names. And again, that's a policy matter, which it is what it is. So there was this mechanism put in place to allow trademarks to be given some recognition in the process. And this is actually a pretty involved protocol, and I'm not going to do complete justice for the protocol, I'm just going to talk about the registry side of this right now.

There are some good documents out there with really fancy ASCII art, which I admire, that you want to look at and hang on the side of your house sometime, it's a nice tapestry. So I'm going to limit the discussion here to what the registry does in these protocols. And it's such that I'm going to talk about two of the phases.

Now when a registry starts work under the new TLD program, 2012 round of applications, there are a bunch of phases out there. And it took me a while myself, as someone who has worked in registries for a long time to understand what these phases are. There is a sunrise phrase. This is one of the opening phases of this saying that, at this point I'm only going to allow people who own trademarks to register trademarks within my system.

There is a trademark claims phase, which says that if anyone registers a name, that matches someone on a certain list, there will be a notice of a trademark claim on that name. It doesn't stop anything, it's just a warning that bad things may come to those who don't like them happening. There are other phases. There is early availability phases, there are friends and family phases, there is presale, and then there is general availability which is what a lot of us are used to seeing here.

But for the purpose of trademark clearinghouse, just a sunrise and the trademark claims periods have any significance. Now this protocol, many protocols involve many parties, it's build over the secured web protocol, again, https. Next slide.

Sunrise is the offering to TLDs of trademark owners only. The registry supplies to the… In sunrise, the registry with tell the TMCH, these are the domain names that are registered out there. The registry receives back a list of trademarks that are no longer listed. So we start out with a list of trademarks that the registry is supposed to protect. And over time, they get taken out as trademark holders are not worried about them anymore.

Basically it's an [inaudible] of names, at this point. It goes from a populated list of things I don't want to register, down to what I'm going to allow. Next slide.

Trademark claims. This is in the early days of the TLD. The registry will tell trademark clearinghouse, the list of names registered that match, they're close matches to the preregistered trademarks in there. Coming back from the trademark clearinghouse, a list of labels comes back saying, here are the ones to watch out for. That's all that's going back and forth at this time, between the registry and the trademark clearinghouse. Next slide.

And again, to show this in blocked diagram. Up on top, the arrow from mountain side to wall slide, I don't go right to left anymore. Says that the service marks, revocation list is sent from the trademark clearinghouse to the registry. During the claims, the middle arrow only the domain name list is sent across that time period, and at all times, the registry is telling the trademark clearinghouse, this is what I've been registering so we can see if we have anything to deal with.

And so, to quickly wrap-up. I don't have a long wrap-up slide, this is basically my wrap-up slide. These are the protocols that we went through. And again, not in any order. EPP was the extensional provisioning protocol for affecting changes in the register of names that are being held in the TLD.

DNS and DNSSEC are the way that the general population of the Internet makes use of what's being registered, knowing where to go with information, for information. The WHOIS protocol is the old way. It has been around forever, saying, for asking who has this name or this address, and let me go find them and knock on the door.

R-DAP is the new way that's going to come along to do the same thing, with the ability to do some more interesting protections of the data. Data escrow is a way to back up your hard drive basically in the TLD sense. And trademark protocols are there to protect the interest of the trademark holders in the registration process of the new TLDs.

So, time for questions. Time for making Steve run to the back.

No, Steve, there is one way in the back first. No here, right here, second row.

I just wanted to make you run.

UNKNOWN SPEAKER: No, I just would like some clarification that the trademark clearinghouse. Is that an optional protocol? Because most registries, I don't think they're using this are they? Especially the smaller ones. I've never heard of that.

ED: I believe that all of the 2012 registries have to use it, have to use it in some way. However, when you say most registries, ccTLDs, I think there is one or two that has gotten, it has actually gotten some whiff of interest in the ccTLDs, because they have the same trademark protections that will eventually hit them. Some of the older ICANN contracted TLDs, I'm not sure if it's filtered to all of the agreements yet, but I'm sure it's being fed into that every time as a renewal. So, but I think of all the new ones, especially all the smaller, if you think the smaller, the new ones are smaller.

There should be some impact in that, in fact, I believe that of all the phases that I was talking about, sunrise and trademark claims are the only two that actually exist in the agreement. Like all of the early availabilities and everything else, their names apply by the TLD for its purposes but are not contractual requirements.

STEVE CONTE: Any other questions?

ED: I was trying to maximize your footwork.

UNKNOWN SPEAKER: Yeah, thank you. I was just wondering about the TMCH. So it is support internationalized script, like fully? Like all the trademark in Thai, in Chinese, and whatever?

ED: Yeah. I believe in the back, way in the back. I mean, a thumbs up for our IETF expert saying it certainly does do that. [Inaudible] it's all the IDNA 2008 internationalized concerns and everywhere else that would get involved here. Yeah. Paul Hoffman, by the way, is our IETF expert.

Any other questions? Someone from over there. No, I'm asking for someone from other there. Please raise their hand. Yesterday's room was a lot bigger, you had further to go.

STEVE CONTE: But you had more people on this.

ED: Yeah, we had more people.

STEVE CONTE: Yeah, this is great.

ED: Yeah.

UNKNOWN SPEAKER: So I'm just asking about the real added value for the trademark and especially for the ccTLDs.

ED: Interesting. I can only give you, off the top of my head type of ideas, but if the ccTLDs, depending on how the ccTLDs it's role in the economy, they may decide that they want to offer the same level of protection to trademark holders in their area of concern that's being asked of in the ICANN area. Some of the ccTLDs, what I've seen over time is a…

I don't want to step on areas I don't want to talk about. As a staff member, I don't want to get into policy and all of that. But EPP as an example. EPP was a protocol that ICANN pushed on the contracted TLDs a long time ago. ccTLDs they were, at the time, going to necessarily follow the ICANN model. But over time, it kind of rubbed off on them.

It seemed like it worked for these registries over here, it made their operations a lot easier, so ccTLDs said, maybe not such a bad idea to do, to adopt that piece of technology. I can see that analogy applying here that the trademark concerns that are being addressed within the ICANN area, maybe something that the ccTLDs realize, maybe we should, maybe we have the same ideas.

That's just my thinking, as I do see over time, an adoption of what's done in the ICANN model by others who are not required to do that.

STEVE CONTE: Just to add to that. I think some ccTLDs use their, some countries use their ccTLDs as a monetarization model. And if we're looking at trademark clearinghouse and their cc is almost entirely to monetize. It's a branding, not going to throw any ccTLD out there, but some of them relate to television, the shortened version of television, things like that.

There might be a way to improve their market by offering some kind of trademark clearinghouse within that. I'm not saying they are and I'm not saying that they're not. I'm just saying that might be one scenario where a ccTLD might embrace and adopt the protocol of trademark clearinghouse.

Questions?

UNKNOWN SPEAKER: Hello. My name is [inaudible]. It's my first ICANN meeting. I miss the part of R-Dap. You said that it has no port number, so how would you…?

ED: Okay. R-Dap is a protocol… The way R-Dap is defined, is it's a way of using the web. And now web servers today have a couple of port numbers. Like the web is on port 80. If I basic http as port 80. Https is on port 443. So R-Dap just uses that. 443 is already open for web access. If you try to do any kind of access on your machine, you'll see that it already has a 443 entry in every firewall, because that's all the secure web.

So R-Dap didn't go off and say get me another port number and I'll connect to that. I'm just going to make a URL, just use the web, and you get to the R-Dap server sitting at that place, on that machine.

UNKNOWN SPEAKER; Thank you.

STEVE CONTE: All right Ed. I'm going to look at it this way. Who is next?

ED: I'm thinking we're almost out of questions. How about online?

STEVE CONTE: Let me double check that.

ED: That's another way to make him run.

This is how I have my fun. I make Steve run around.

STEVE CONTE: No, I do have Patrick Miles typing though. So we're going to wait in suspense for just a moment. No, he stopped typing. Excellent presentation Ed.

ED: Thank you. I owe you another dollar, Patrick I mean.

STEVE CONTE: All right. Well Ed, I want to thank you for taking the time today and yesterday, for presenting on the registry protocols. We do have, let's take a moment and thank Ed for a second.

We do have another session coming up in about 45 minutes with the root server operators. This is from our root server advisory, security advisory committee, RSSAC. One of the co-chairs, Johan Liman will be presenting about the root part of the DNS, and also what the RSSAC is doing lately within the ICANN structure.

Great turn out yesterday, it was a great presentation, super interesting. So that's at 3:45. And then not our presentation, but after RSSAC will be a beginner's guide to DNSSEC in the same room. And if you're just getting into DNSSEC, I highly recommend sticking around that.

I am going to beg and plead. We are looking for active, actively looking for feedback. So if you have a moment before you pack up your laptop, please head over to this URL. We're looking for how well we're doing. We want to keep doing presentations and tutorials like this to bring technology to the community here to help understand what the technology does, and help inform the dialogue during the week of discussions.

We also working to avoid the diminishing return model, so we don't want to keep doing the same presentation ICANN meeting after ICANN meeting. So any ideas in what we could do, what you as a community member would like to see explained, or discussed from a technology perspective, we would love to hear from you.

So before you pack up and run off to whatever meeting you have next, please just take a moment and fill that out, and we would appreciate your input. With that, we will be back on in about 45 minutes. So you have some time. Thanks.

[END OF TRANSCRIPTION]