DUBLIN – How It Works: Internet Networking
Sunday, October 18, 2015 – 11:15 to 12:45 IST
ICANN54 | Dublin, Ireland

DAVID CONRAD: We're just waiting for this presentation to upload, so as soon as it's up on the screen we'll get started. Welcome, everyone. We'll be starting up in just a little bit.

Hello, everyone. I'm David Conrad, ICANN CTO. If you're just joining us, we did have a previous session on Internet Standards Setting. This session is on Internet Networking. It will be presented by Alain Durand. This is the second of the series of tutorials. The first session of these was provided at Buenos Aires. People seemed to find them useful, so we decided to continue this. If you find these tutorials useful, you will be given a URL for a survey, and we'd very much appreciate any feedback that you can provide on how we did in these tutorials, what you'd like to see in future tutorials. That URL will be provided at the end of this session.

Right now we're just setting things up to allow the presentation that Alain will be providing to be recorded using the Adobe Connect stuff. And I think we're just about ready, so if you give us a couple more minutes or a couple more seconds... Seconds.

We're down to seconds. Thank you. I will now throw it over to Alain.

ALAIN DURAND: Good morning. My name is Alain Durand. My work address is 801 17th Street, NW, Washington, DC, USA. And I came here through direct flight from Washington Dulles to Dublin Airport.

I was just talking about my name, my address, and the route I took to come here. This presentation is about just that. It's about how things work on the Internet, how they use names, how they use addresses, and how they use routes.

I gave essentially the same presentation at a previous ICANN meeting in Buenos Aires, and we're simply going to go through it. I would like to make it as interactive as possible, so if you have any questions, please jump in. Try to have opposite style from what Paul made, and we'll see what works better.

The agenda today is we're going to talk essentially about four different things. First one is networking by number. We may remember the ISO model, the seven layers of ISO model, and I will actually extend it to eight and nine and we'll see how it works. We're going to talk about naming, addressing, and routing, as I mentioned earlier.

Networking by numbers, layer one through layer nine. Actually, I'm going to start from layer zero. Layer zero is a physical layer. When you talk about networking, you really have essentially two types of layer zeros or physical networks, either wired or wireless. They have very, very different properties, and some of the usage of it depends on those properties.

Why do we have either copper networks or fiber networks, wireless when we have all those antennas? Typically you'll get more bandwidth out of a wired network, but one network is essentially from one point to another point. If you want to have an area of coverage, wireless networks are more interesting.

If you deploy a network in a place where there is nothing, well, wireless is much easier. You very quickly cover a very large area. If you want to bring wired to every house there is, you have to dig some trenches. Like in the picture over here, it can be expensive, it can take time, get a lot of authorization before you do that.

An antenna, you have to put an antenna in and you're done. However, in some neighborhoods, like in some residential neighborhoods, people don't like to have antennas next to the house, so that's not necessarily the best solution there. Then you may have to rely again on wired and some copper. So the two have different properties. I'm not going to go into spectrums or anything like that, but there's some interesting discussions about how you can share those medias.

What essentially comes through layer one? When you have a fiber, for example, you are not the only one using this fiber. What you do is you send wavelength (essentially light) at a specific color like a green, a red, a blue, or different shade of that. Each color has a different wavelength, and you can have different wavelengths in the same fiber. So in the same physical fiber you can have multiple persons sending different types of data.

When you have this fiber, you send data from the beginning of a fiber to the end of a fiber. So if you would like to have a communication from, let's say, Washington, DC, to here, if you want to have just one fiber, you will have to create your own fiber over there and it will be a very, very long fiber. It might be difficult every time you want a new communication to layout a new fiber, so people don't do it.

What people do is they create a layer two network by assembling a bunch of those different fibers that already exist so that they create what they call a fiber path. They will take a wavelength in one fiber from, let's say, my building to a central office near where we are, near the White House, where it would have a certain wavelength.

Wherever we connect that with a switch to another fiber that may go from there to some place near the shore, and then they connect that to another fiber that goes over the Atlantic, and that may go through different hubs, and all of this with different colors. Potentially, you can transmit in blue in one side, and red in one side, and green the other, but all of this may be patched together to create, essentially, a fiber path between, let's say, my office and here. So that will create a layer two network.

Examples of layer two networks may be an Ethernet cable that goes to different laptops. We used to go into different laptops many years ago. Now, we use 802.11 Wi-Fi networks to do that, so that’s another example of a layer two network that connects multiple devices.

This creates one network. In a small network like this or a big network like this, [the property] is that every device will see every other device. So if you send a message to "all the devices on my network," everybody will see that message. So a potential for [storm servers] messages.

From a personal perspective, that's not exactly desirable. You want [some isolate] those networks together. If somebody creates, for example, a [storm] in this room, it's not going to [be seen in] another room next to here, or it's not going to be seen in another country or another continent. So we create topology to isolate for different networks. And layer three is essentially what this is about, connecting or interconnecting those different networks that we have built [recently].

IP protocol for Internet protocol enables to connect all those layer two things together and essentially provide the addresses that are global to all of us.

The moving of the stack transport. Now, you have a network connection, and you want to be able to send packets, otherwise known as datagram. There are two main protocols that have been defined by the IETF. One is TCP; one is UDP. UDP is the simplest one. You send a packet — that's all it is — or you receive a packet.

TCP tries to answer a different question: "I want to make sure that I sent some data, but I want to make sure that data has been received." So I want to send something to Steve. Hi, Steve.

STEVE CONTE: Hi, Alain.

ALAIN DURAND: Hi. I receive your 'hi.'" Now, I know that Steve knows me, he has heard me, because he sent me back some information, and I'm sending another message to him to say I've acknowledged the fact that now you hear me. This is what's called the freeway handshake in TCP. Then we send data. I'm going to send to Steve, "Steve, this is my slide. Did you receive it?"

STEVE CONTE: Ack.

ALAIN DURANT: This is my second slide. Did you receive it?

STEVE CONTE: Ack.

ALAIN DURANT: There you go. That’s how it works. This is my slide #99. Did you receive it?"

STEVE CONTE: I didn’t get that one.

ALAIN DURANT: I can’t hear that. Did you receive it?

STEVE CONTE: Is that an "Ack"? "NAck"?

ALAIN DURANT: That's a "No Ack." That's a bad [parse].

STEVE CONTE: No Ack.

ALAIN DURANT: So we have to hear that one better. So if I don't hear it or he sends me something that I can't understand, it means we're having something wrong in the network. What am I going to do? I'm going to send it again. Steve, here's slide #99.

STEVE CONTE: Ack.

ALAIN DURANT: Now I know that he got it. That's how TCP works: making sure that the party on the other end really heard my message. And I'm going to transmit again and again and again until he hears it.

Sometimes I'm going to be a little smarter. I may not send the entire page of slide #99, because it's too big or I'm sending it too fast. So we have what we call TCP windows, and I'm going to essentially reduce that window. And it says, "Maybe I'm going to send… This is the first part of page 99. Can you receive it?"

STEVE CONTE: Ack.

ALAIN DURANT: There you go. Now I can send the second part. And like that, I can adapt to the bandwidth that is between me and Steve. If there's lots of bandwidth, I can send everything. If there is congestion, there are other people talking at the same time, maybe I will back off for a while and then start again.

Here's the part of page 99, Steve. Did you receive it? Oh, not working. Maybe this is still too fast. I'm going to break the second page again into two subpages. This is page number 2.5. Did you receive it?

STEVE CONTE: Ack.

ALAIN DURANT: There you go. Now that you've received it or maybe there's no more congestion, I can try to go faster this time. Here are the next three pages. Did you receive them?

STEVE CONTE: Ack, ack, ack, ack.

ALAIN DURANT: There you go. That’s how it works, and it's really adaptive. It can slow down. It can go faster. As a user, you don't have to do anything. It's all taken care of by the protocol underneath. Both TCP and UDP have some other protocols that are mostly used for specific applications. Ones that are the most known today are RTSP, for real-time streaming protocol. That's when I want to stream a movie.

When I want to stream a movie, this is very different from making sure that the file is completely transferred because if you think about it, I'm transferring the movie. I don't want to pause and then go back 30 seconds later or 30 seconds before because there was a packet that was lost and retransmit what happened in the dialogue a minute ago. That makes no sense. I'd rather like to drop a few frames but keep the sequence of time intact.

So those protocols have different characteristics. I want to really make sure that we can keep the flow running, and the integrity of the data is not as important. That’s why we have those different protocols here.

Now, all this is just data. I'm sending binary data. How to make sense of this data? That's what layer six is about. It's about presentation of this data.

So it could be like anything goes, like just [pure] binary and then it's up to you to make sense of it. It could be some text files, and there are different versions of that have been defined over time. The most recent one is a format called JSON. It's essentially a dictionary of things. It's really good to create, for example, menus or to create actions. This is used to describe the configuration for network equipment, for example.

This is an example here that creates a menu. You have an ID. You have some values. This is a pop-up. You can describe — and there's a semantic that is embedded now in [to this] language, so it's [helping] that application will have an easier time to actually pass.

Of course, there needs to be a standard, because if I describe something in the language that Steve doesn't know, he's going to look at me and say, "It doesn't work. I don't get it." That's why we have things like standards bodies like IETF and a couple others so that Steve and I can agree that this is the standard that we are going to follow. And that's the value of standards is that now we can communicate.

Both layer six and layer seven, which are the ones you are probably the most familiar with, it's called applications. Now we can have almost anybody, regardless of age, using the Internet. We have kids now that are much more proficient in using the Internet, browsing it, than I am, and I've been doing that for way too many years, so maybe it's just that I'm getting too old. The most used protocol on the Internet [inaudible] web.

Paul was talking about the web being defined a while ago. The protocol is actually HTTP or HTTPS for the secure version of it. Obviously, [inaudible] organized at a higher level is the work of a W3C.

Now, everything (mostly) on the Internet is one version or another of HTTP. Above that, we have what we jokingly call the financial layer. None of this makes any sense if you can't finance it. Nobody's going to lay out any fiber if they can't have any return of investment. Nobody's going to create a new application if they cannot find users to go and deploy those things.

So making sure that there's some kind of a business case and or there's some kind of usage behind it is really important. It's not just about money here. It says financial, but it's really more about: are people going to really want to use this thing? If you are not solving a problem, but it's interesting, so maybe it's not really that important.

And on top of that, there's all the political layers. There's a different relationship between standard bodies with different relationships between, for example, IETF and ICANN, ITU and all the others. All this is part of this huge hourglass, what we call a protocol stack.

So here we are now in this galaxy somewhere far away. This is all the different layers that you need to somehow think about talking about networking.

Now, this is kind of a true story. One day I started to have a really bad toothache. I don't know what I had for dinner that day. It was maybe a piece of rock in the food that broke one of my teeth. Ask Steve, "Who is your dentist?"

STEVE CONTE: Well, you know, Alain, here in Dublin, Dr. Guinness is my favorite dentist.

ALAIN DURANT: Let's go and find this Dr. Guinness. I really need some help here.

What is a name? Look at the dictionary. Name is a noun. It's a word or set of words by which a person, an animal, a place, or a thing is known, addressed, or referred to. Example: My name is John Parson, name is [inaudible] is in Germany, so German name is for [Cologne]. Or it could be another definition is a famous person, like a big race, we love to talk names. Essentially, it means when I know your name, I know who you are. I know Steve's name, Steve Conte. I know his name. I know who he is. I can talk to him.

So Dr. Guinness. Having a name is not useful in itself. It's only useful when you have a communication with somebody. It's a way for us to relate about a certain person. I can talk to David. David, do you know this Steve guy?

DAVID CONRAD: No.

ALAIN DURANT: Yeah, maybe I should think twice about this recommendation for Dr. Guinness. Maybe Dr. Guinness is not really a solution here.

[DAVID CONRAD]: Maybe Dr. Jameson would be better.

ALAIN DURANT: Probably. So we can talk to someone, like I just talked to David, or we can talk about someone. I talk to David about Steve. [inaudible] things that you can do with a name. And, of course, we can pass it on from one person to another, like Steve gave me a referral to your dentist. What's again the name of your dentist now?

STEVE CONTE: We're going to go with Dr. Jameson.

ALAIN DURANT: Dr. Jameson. Okay, maybe Dr. Guinness was not really that strong.

Names have scopes, and names, some are unique within a scope. In my family, I'm the only one called “Alan” or Alain, depending if your French pronunciation or American pronunciation. Nobody else is called Alain. So when a cousin of mine says, "Oh, Alain now lives in America," everybody knows that they're talking about me.

But when I was in grade school, Alain was a fairly common name. There usually were three or four students in my classroom who were called Alain, so when the teacher said, "Oh, Alain, can you please come to the blackboard?" three of us were looking at each other and said, "Which one?" Not good. So it was ambiguous, and that's a bit of a problem.

When the teacher says, "Alain Durand, please go to the blackboard," then I knew it was me. I knew what to do. The others could just stay in their chairs and do whatever they wanted.

But a name is not really enough to communicate. Okay, I know I need to go see Dr. Jameson, but where is Dr. Jameson? I have no idea. So what I need to do is to map the name to some kind of another identifier, lower level identifier in the stack, that will enable this communication. This process is called name resolution.

What I will do is I will go in some kind of a directory, and I will look at the letter J for Jameson, and I will try to find Dr. Jameson in there. I can use a phone book. I can use the Internet to do that. I can use all kinds of other things, but in the end I would discover that Dr. Jameson is on 125 Root Canal Road in DC, or in Dublin, wherever you want. We are going to try to go to 125 Root Canal Road.

So just to conclude the path names, I want to briefly mention some of the recent issues that have been discussed the IETF context or the ICANN context. The first one is internationalization. Initially the names were ASCII character strings. That's fine if you are English speaking. If you are French speaking, you start to have problems because we have characters with E with an accent or C with a cedilla. If you are Chinese or if you are Japanese, that's much, much worse, because characters don't even look like this. So a lot of people spend a fair amount of time trying to fix this problem.

The second one was authentication. I talked about this directory to map the name to an address, like Dr. Jameson is on 125 Root Canal Rd. Well, I want to make sure that the directory gives me the right information, because if it's a fake directory – maybe there's a conspiracy of doctors somewhere who are trying to make sure that I go see Dr. Guinness and not Dr. Jameson, so they're going to give me a fake address, the address of Dr. Guinness and I don't want that. So how do I know that the directory I have has the right information? All this is done through a technology called DNSSEC. A lot of cryptography See Paul or Ed, he is a specialist of that, if you want to have more information about it.

The third topic that got a lot of coverage here at ICANN is the extension of the root zone, the top-level domains. It used to be only a country code top-level domain, like dot-uk, dot-fr, dot-jp, dot-ie, and some generic top-level domains like dot-com, dot-net, dot-org, dot-int, dot-edu, dot-mil. But now we have all these new gTLD programs, and that has been extended quite a lot. So this is still a discussion happening within ICANN. I'm not going to cover it too much here.

Where is this Dr. Jameson? We know it's 125 Root Canal Road, okay, but let's dig into this more. What is an address? Back to the dictionary. An address is the particulars of a place where someone lives or where an organization is situated. So if you know your address, you know where you are. Remember, if you know the name, you know who that person is. If you know the address, you know where that person is.

Let's make a little detour. So I told you I live in Washington, DC, and our building is next to a very famous smaller white building that we like to look at. So let's talk about it a little bit. The address of that famous place in DC is 1600 Pennsylvania Avenue, NW, Washington, DC, 20500-0003, USA. That's the complete address of this place.

If you look at it, there's a hierarchy. Start by the rightmost, USA, so that's a country. Then DC for District of Columbia. That's the state, although DC is not exactly a state. NW is because DC has been separated into four quadrants. There's northwest, southwest, northeast, southeast, so it's the northwest quadrant.

Pennsylvania Avenue is a name of a street or the avenue in that case, and 1600 is simply the number within that street.

Names are not always organized on a geographical way like this. Sometimes just a flat structure. So if you live in America, you may have the phone system. When you want to make a toll-free call, it's a 1-800-number. You have no idea where the caller is. Actually, he may not even be in the US. He may be rerouted somewhere in an island, somewhere in India, somewhere in France, anywhere. It doesn't really matter. There's no notion of geography whatsoever in there.

Cell phone numbers. For example, in France, all the cell phone numbers start with 06, but I have no idea where in France that phone is going to be. And actually that phone might be roaming in another country in Europe, and I still don't know where it is. All I know is it starts with 06.

Same thing with IP addresses. There's nothing in the structure of an IP address that tells me where this IP address is. Now, there's a database that has been built over time by people who operate the network that know that this IP address has been allocated to this particular service provider, and the choices that have been done, and we can geo-locate with some specific database to trace this address to one particular place. But this is done after the fact. There's absolutely nothing to date in an IP address that says, "This belongs to France or to the UK or to Ireland or to Germany or to the USA or China" or anywhere else. That's just been recorded after.

Similarly to names, addresses have scope. So if you live in DC and you want to go to the White House, all you need to say is, "1600 Pennsylvania Avenue, Northwest." You don't need to specify the full address because you're already in DC. People know you are there. The northwest part is important, because there's another Pennsylvania Avenue in the south part of the city, so you need to specify that, but that's about it.

Similarly, if you live in Europe and somebody's talking about Paris, most probably you're talking about Paris, France. But if you live in America, there are 29 cities which are named Paris. If I say, "I live in Paris," there's one very, very little village — maybe have, I don't know, 50 people living there — that is about 20 miles from where I live that's called Paris. Sometimes I go to Paris for the day. I don't have to take a plane. So this is ambiguous. Same thing with names.

Also similar to names, an address is a handle. I can use it directly. For example, I can send a postcard, and I write the address. I write inside of a postcard, or I can just give somebody else the address as a reference. For example, instead of going to the directory, I could have gone to Steve.

Steve, I want to really go to see Dr. Jameson. Do you know where he lives?

UNIDENTIFIED MALE: He's at 123 Root Canal Road.

ALAIN DURAND: Do I really trust this guy? Do I have a cryptographic signature? Because I've seen in another directory that it was 125. Maybe 123 is really Mr. Guinness. There's a conspiracy of dentists around here.

So once again, this is why we need to have some cryptography, and we need to have some signatures about information we get from directory or when we get information from someone. Maybe the channel of communication between Steve and me was corrupted and somebody injected the wrong number in my audio. That's a possibility, or maybe Steve's database was corrupted. So simply securing the channel between Steve and me is not enough. Maybe I need to go back and ask Steve, "Did you check your computer? Did you get a virus last week?"

UNIDENTIFIED MALE: My database is often corrupted, so…

ALAIN DURAND: So those names, those addresses can all be passed to somebody else as a reference. That’s the part that’s important here.

Alright. So when you have an address, is that enough? Okay, going back to my example of the White House, I'm going to write the address of the White House on a postcard, and I want to send this to the US president. If I simply write the address on the postcard. That doesn't make that the postcard will arrive to the president. Somebody has to go and deliver the postcard. If it was not for the post office system, I could have a paper and that's just about what it is. It will end up in a bag.

Because there is a system that is going to take this and move it along to the final destination, then I know that my postcard will arrive. And if I trust the system, I know for sure it will arrive. Usually the post system has been pretty reliable. So there are some cooperative agreements between different post offices in different countries, so if I put a postcard here in the mailbox, in a post office box, outgoing post office box in Dublin, most likely it will arrive in DC. This will go probably through a number of hubs, but it will arrive eventually.

So let's look at Internet and Internet addresses. We have two protocols that have been defined on the Internet. The one that most people are using today is called IPv4. It was defined in 1981, and is used massively till now. The new one is called IPv6. It was defined in 1994 and is being deployed. It has been being deployed for the last 20-something years.

The difference between the two: IPv4 is a 32-bit address field, so that is about 4.9 million addresses. Many of them are reserved for multicast addresses and other specific pools of addresses. In practice, you have 3.2 billion addresses usable. The population on Earth is about six, seven billion people, so that's clearly not enough even if you just wanted to have one address per person. Not even talking about giving an address to [inaudible] or giving an address to a light bulb or anything. This is simply not enough.

Question?

UNIDENTIFIED MALE: Alain, a quick question. What is the significance of the numbers four and six, I mean, why is it called IPv6?

ALAIN DURAND: Why is it not called IPv5?

UNIDENTIFIED MALE: Yes.

ALAIN DURAND: Okay. So when IPv next generation was defined and was written down into an RFC, it went through the process that Paul just explained an hour ago, somebody went to one of the IANA registries and looked at what will be the next version that will be available in the registry. It turned out that somebody had registered version five already. It was for a protocol called SD-1 or SD-2 that was for streaming data protocol. So they said, "Okay, there's already a version five registered in this IANA registry, so we're going to take the next available version." So they took IP version six. That's the value of a registry.

UNIDENTIFIED MALE: So you're saying it has got nothing to do with the number of bits?

ALAIN DURAND: No, absolutely nothing to do with the number of bits. Actually, there were a number of candidates in IP next generation to define what is now IPv6, and each of them was assigned a number. If you go to the IANA registries that define their IP protocol version, you will find IPv7, IPv8, IPv9. There's a bunch of them that have been reserved. So the next one, actually, is probably going to be, if ever, something like IPv11. We have to make sure that we don't reserve too many of them, because it's a 4-bit field, so we can go to 15, right? After that, we may have some issues.

IPv6 is 128-bit addresses. This is not simply four times more than 32-bit. If you look at the number of addresses and do the math, this is this gigantic number down on the screen there. It’s all powers of twos. So I don't even know how to pronounce this number.

We did some math back in the days when we defined IPv6, and there's enough addresses to put an identifier to every single grain of sand on the planet, so we think that we should be good for at least another few years, another 10, 15, 20, 50 years. After that, well, maybe my kids or my grandkids can find another solution.

So IPv4 addresses. As I said, there are like 3.2 billion of them usable, and they have, for the most part, all of them been allocated. There are five registries. The first one that got created is RIPE in Europe. There’s APNIC in Asia-Pacific. There's ARIN in North America. There's LACNIC in the Latin American Caribbean, and AfriNIC for Africa. AfriNIC right now is the only one who has a few addresses left, very few. ARIN was the last of the big ones to run out. They officially sent a few weeks ago a notice saying, "Out. No more."

RIPE and APNIC had a different policy put in place to say for the last block they were not going to normal allocation, but if you are a new entity, you just show up, and you will be allocated a small block and that's it. You have one allocation. No more.

For all intents and practices, you cannot get any new addresses from the registry today. But the Internet is still growing, so what do people do?

Well, sometimes IPv6 has been positioned as a replacement for IPv4. The problem is they are not compatible. This is the same as the left-hand side of the picture, this is a US plug. On the right-hand side of the picture, this is one of the Asian plugs. I'm not sure if it is from Australia. You cannot just force it. It doesn't work. That's a technical limitation that has nothing to do with policy. It's just the way it is. On one hand, you have a 32-bit space. On the other hand you have a 128-bit space. It just doesn't match.

So as a result, you cannot simply say, "Oh, from now I'm simply going to do IPv6," because you need to take care of all the things that already exist. If you were to design a brand new network, yes, but that's not the case. So not every equipment on the Internet supports IPv4 and IPv6. My laptop does. Steve, does your laptop have IPv4 and IPv6?

STEVE CONTE: Yes, it does, but I actually had to turn my v6 off because my v6 at home doesn't work very well.

ALAIN DURAND: That's a very, very typical situation, unfortunately. Things kind of half work, [half there], but you cannot really rely on them until we are there 100%.

So you say that you have a home router, Steve, at home, right? It does not support IPv6. That's what you are telling me?

STEVE CONTE: It does. I'm not sure if they're doing tunneling or they're doing native v6, but it's causing massive latency issues at my house, and then it reverts v4.

ALAIN DURAND: So if the service provider were to do the right things and bring native IPv6 to your home and give you a new router, you will have v6 in your home, right?

STEVE CONTE: Theoretically, yes.

ALAIN DURAND: Sounds like the problem solved, right? Maybe not. Do you have a TV at home?

STEVE CONTE: Yeah.

ALAIN DURAND: Is your TV streaming things?

STEVE CONTE: Yep.

ALAIN DURAND: When did you buy your TV, a couple of years ago?

STEVE CONTE: No, couple months ago.

ALAIN DURAND: Oh. Does it support IPv6?

STEVE CONTE: I honestly don't know.

ALAIN DURAND: Well, I checked on my TV that I bought maybe just two years ago, and the answer was no. The firmware is an IPv4-only firmware. So this TV, for the rest of its life, is going to do IPv4, never do IPv6. Even if my service provider does the right thing, even if I change home gateway and all those things, it will still be an IPv4-only device.

So if my service provider wants to offer me a service that I can use, they cannot give me an IPv6-only service. It has to give me both. That's a problem, because now you still need to maintain an IPv4 service. So how do you do that? Well, essentially, the way you do it is you share resources. There's not enough for everybody, so we have to share.

I was at a pub last night, and the pub was fairly full. I had to share a table with somebody else that I did not know. Was my food any worse or any better? No, it's the same food. I even had a conversation with those guys. That was fine.

On the Internet, that's the same thing. We have to share IPv4 addresses. The same way if you have a home gateway you share. If there's only one address, and you share it among all the devices that are in your home. If you have [inaudible], you share the address with other people that are in your neighborhood. That's what this is about.

So how many people can you put behind the same IP address? Some of the studies that I have done in my past lives have shown that ten people, piece of a cake. One hundred people, very, very easy. A thousand people? I've seen that being done. If you multiply simply by 100 and you get a block of addresses — let's say a /16 IPv4 block, which is 64,000 addresses, multiplied by 100. That's six million. For a relatively small block, now you can serve a fairly large population.

So how do we get those small blocks? You get them today from the transfer market. So all of the RIRs — actually, LACNIC is in the process of [defining] it, but all the other ones — have now some policies in place that allow for transfer of resources, and there's a marketplace associated with that. Some terms and conditions may vary. The ARIN vision right now you still need to demonstrate that you need those addresses to be registered to your name. In other places, you don't have to demonstrate that need. But for all intents and purposes now, what happens is you transfer addresses.

For example, there are folks who have been allocated a large set of addresses in the past, and they may or may not use them all, and they can simply transfer – usually it involves some exchange of money, but it doesn't have to – somebody else.

Again, if you are a new service provider and you want to say, "I need to offer a new service," you're going to do IPv6. You must maintain IPv4, so you will get transferred one of those blocks. Let's say it's /16. You use the [NAT] technology. You multiply the space by 100. So /16, 65,000 addresses multiplied by 100, six million. Yeah, you are good to go for a while.

Those are some statistics of how many addresses have been transferred. [ARIN won't start until May]. I guess I need to update those. But this is the number of transfers, so you see there are about, in the three main regions (RIPE, ARIN, and APNIC) and you see that there are about 400 transfers per month. But all transfers are not equal. Some are for small blocks. Some are for large blocks. So if you try to normalize this in terms of how many /24s equivalent, we see that at the peak in February, there were 25,000 /24 equivalent. That's about six million IP addresses. So that's a significant market. That's the least we can say.

So going back to my dental problem. I really need to go see Mr. Jameson. I have his address. I'm pretty sure it's 125 Root Canal Street, not [123] as some people wanted me to believe. So how do I go there? I need a route to go there.

A route, going back to the dictionary, is a way or a course taken in getting from a starting point to a destination. So if I am here in this building, I need to go to 125 Root Canal Road. Obviously, I'm not going to take the same route as if I were in my office in Washington, DC, going to the same place. I start from a different place, I go for a different route.

If you remember, a name is who you are. An IP address is where you are. A route is how to go there. If you remember only one thing from my tutorial, this is it.

In order to build routes, the way it’s done is by building them before traffic flows. It’s like if you drive a car, you will see signs on the road say, "Cardiff this way; Dublin this way; London this other way; New York a different way." Somebody has put those signs in place before you arrive. It’s not just popping up instantly.

Same thing with the Internet. Somebody needs to build those routes before you want to send traffic. How do you build those routes? Well, it's done dynamically. It's done by reversing the flow. The destination is going to advertise where it is.

On the right-hand side is a destination. It's going to send messages to its neighbor, saying, "You can reach me directly through this link. I'm here. I have this block of addresses. If you want to send a packet to this block of addresses, send it to me."

The yellow node on the right-hand side is going to re-advertise this message to his neighbors, and say, "Oh, if you are interested in going to this block of addresses, I can get there. Send your packet to me."

Similarly [inaudible] nodes, we are now one step away from the [rightmost one], are going to send the same message to the other neighbors, saying, "I know a guy who can send traffic to the final destination." And you do that recursively, and it becomes "I know a guy who knows a guy who knows how to get there."

This is a cooperative system where all the nodes need to cooperate to listen to those announcements and say, "I'm going to hear them, register them, and re-advertise them to my other neighbors."

When I'm going to send the packet, now I'm like the car arriving at the intersection. I see all those signs in the road. That's exactly the same thing. I'm sending my packets, and I'm going to look up in the table. It says, "I have three exit networks. Where do I go in order to reach my destination?" And it says, "Go to the bottom link and go to the yellow router on the bottom."

Yellow router is going to do the exact same thing. It's going to look at this table. It says, "I remember I knew a guy who knew a guy who knew how to send this packet there." So he's going to send it to another router that is closer to the destination.

This process will be repeated a number of times until you arrive to the last router that says, "Oh, yes, I know exactly where it is. It's over there."

Again, this is a cooperative system. We need to rely on the other service provider to do the right thing. This is where it goes back to the nature of the Internet, to what Paul said earlier on. People cooperate well. They don't simply say, "This is my network," or "This is your network. No, we have one network, and we all participate in it. We all cooperate."

All the service providers in the world do cooperate to make sure that this thing works. There is not an Internet police that says, "You need to cooperate or I'm going to make sure that this packet goes there." No. It's simply based on cooperation.

Now, if you remember in my previous discussion with Steve about what [inaudible] address of Dr. Jameson, we have the same issues with routing. I remember I was telling you somebody says, "I know a guy who knows a guy." But if there's a bad actor in the system or somebody who has been corrupted or compromised and injects some wrong information. Like, for example, there's a bad guy that speaks really loud and tries to convince somebody else in the path that he owns the address, the block of addresses, for this particular network. Well, if you insist, then somebody’s going to believe him. Just receiving messages. That’s how you get traffic redirection to some bad places.

In order to prevent from that, there is a technology called RPKI for Resource Public Key Infrastructure that is going to apply some cryptography and signatures. The same way that we have DNSSEC signatures on the names, but we had also IPSEC that can secure connection between two end points.

Here we have RPKI that has a cryptographic signature of the origin of the announcement. So if a guy, the router that is in the node to the right and the bottom here, that receives the announcement from the bad guy saying, "Please forward the traffic for this destination to me," he's going to look at the signature of this announcement and see it doesn’t match, so that's obviously a bogus so it will simply drop it. So it will not listen to the bad guy because the signature is not there.

This RPKI is a system that has been discussed for a long time. It's not in full production yet. There are numbers of issues. Is it going to a centralized or a decentralized system? Is it only to provide the validation of the origin, or is it going to validate also all the other routers in the path? All those are still open items for discussion, and that's why the technology has not really matured to a point where it is deployed everywhere.

I had a name. Remember, Steve told me. The first time was not a good try, but the second time, Dr. Jameson. The address. What's the correct address again?

STEVE CONTE: 123 Root Canal Road.

ALAIN DURAND: Wrong address again! 125.

STEVE CONTE: 125? Darn.

ALAIN DURAND: See? And now I've found an actual route to take me there. I eventually meet Dr. Jameson who can take care of this issue with my teeth.

So that essentially concludes my presentation. I will be happy to answer any other questions.

UNIDENTIFIED MALE: We do have a comment from remote participation, if I could read that. This comment comes from Michael Oghia. I apologize if I mispronounced that. He's the 2015 ISOC IGF Ambassador. His comment is… This was when we were talking about IPv4 and 6. "Since we're discussing IPv4 and IPv6, we're putting a lot of work into the IPv6 IGF Best Practice Forum. The working document is currently open for comments, edits, and suggestions. If this is an area of interest to you or to the community here, please feel free to comment." And then he gives the URL, which is in the Adobe chat room, and you can go to the history on this and get that URL.

ALAIN DURAND: Thank you very much. I can't stress enough how important all this work on deploying IPv6 is. Just keep in mind, this is going to be a long process. It's not going to happen overnight.

Another question?

UNIDENTIFIED MALE: Just talking about IPv6 and what the business user should do, I'm thinking in the environment of a business, say, with a website a hosting contract with a hosting provider. Where are we with that? Are we on IPv4, or we don't really know, do we? Do we leave that to other people? Do we need to think about it? Thank you.

ALAIN DURAND: If you are in the business of hosting, for example?

UNIDENTIFIED MALE: Not in the business of hosting, but as a consumer of hosting. This a website with a contract hosting somewhere.

ALAIN DURAND: Yes, that's an interesting question because I talk about the ISP case and ISP needing to support both. Now, it's interesting you talk about the user case, what a user needs to do. At this point in time, as a user the message has always been the IPv6 thing is not real.

If you look at some of the statistics of traffic, you see that for many, many years IPv6 [wasn't even noise]. Zero percent, 0.1%, barely one percent. If you look at the more recent things that have been happening in the last 12 months, for example, in North America — there was a very good presentation from Jeff Houston at the NANOG meeting just about a couple of months ago. He was showing that you get now a number of requests at Google, something like 8% or 10%.

If you look at the number of users that are reachable for IPv6 in the USA, it could be about 40% or 60%. If you look at really the number of users who get it and the number of devices that can get it, maybe it drops to 10% or 20%. But 10% is a lot, especially when we were at 1% only a year ago.

There is now a real movement to get more and more IPv6. This is real. This is not just a hypothetical anymore. So now it becomes really interesting for content providers or users who have their own things on the Internet to say, "I need to hop on this bandwagon. I need to become IPv4 reachable and IPv6 reachable."

Being IPv6-reachable only probably is a little too early. But making sure that both now are being enabled becomes important because you see a lot now of, for example, of wireless devices and wireless network but go for IPv6 and IPv4 they start to get for different stages of translators tunneling whatever it is.

If I look at my looking glass, I can I can see that the IPv6 path is going to become more and more qualitatively better quality than the IPv4 path. The IPv4 is going to degrade. The IPv6 is going to get better. If I want to better serve my users, now is a good time to start serving them with IPv6 and ask those hosting places to deliver IPv6. Ask those registrars who have my names in the DNS to make sure that they can also put an IPv6 address in there. Ask the registries who [the] registrar is a customer of to register this properly with also IPv6.

DAVID CONRAD: There's also increasingly services, particularly if all you're doing is a website, some content delivery networks are providing sort of a gateway service so that if you use their content delivery network, your website will show up both with IPv4 and IPv6, even though your back-end, the actual content itself, is only available via v4.

Not to promote any particular one, but I do know that CloudFlare is one of these content delivery networks that provides this particular service. I don't know if there are others at this point in time, but I suspect there will be because it makes moving to IPv6… You as the content owner, the person who runs the website, you don't have to do anything. The content delivery network provides that for you.

ALAIN DURAND: Actually, I can answer your question, David. I know that Akamai and [Limelight] have had this for a while too. The technology we were talking about, the [NAT] technologies, we can use also some reverse-[NAT] technologies to put in front of your server. So let's say that you have an infrastructure that deploys IPv4 servers, and you want to use this infrastructure to also serve IPv6 users. Well, you put the [NAT] in front of it that will take the incoming v6 packet and translate it into a v4 request to your system. And those technologies are available from many vendors, and they work actually quite well.

Next question?

UNIDENTIFIED MALE: So I've heard that there are various routing schemes available, one of which is anycast. Could you please shed some light on the anycast routing technology?

ALAIN DURAND: Everything I've talked about here is about unicast. Unicast means you have one source and just one destination. There are other technologies like multicast and anycast that exist. Multicast is essentially used for group communication. It was initially done for… You are part of a group, and you send a message to everybody in the group. Everybody receives that. Later on, it was refined as a single source, which is really used when you do video distribution. You have one video source, and you want to send to many different recipients. Anycast is a different technology.

Anycast means I want to send to somebody, and I know that there might be multiple persons, but it doesn't really matter. I want to send to one person.

Think of it as you are in the supermarket and you're doing your grocery shopping and when you go, you have to pay. There's a line of different registers. You can go to any of the registers; it doesn't matter. Functionally, they are exactly the same. If one register gets closed in front of you because a person is on his lunch break or whatever, you simply go to the next register. Anycast is exactly that.

It enables you to send a packet, and you don't really know who's going to handle it. It doesn't matter. This is used a lot for things like DNS resolvers. So you will send a packet to a DNS resolver, and being one resolver or another or a third one from the same company, it doesn't really matter. You want to make sure that the packet is sent to the closest to you or the one that's the least busy or whatever policy has been put in place. The reason it works is that they will all provide exactly the same service. So asking Resolver 1 or Resolver 2 or Resolver 3, it doesn't really matter.

UNIDENTIFIED MALE: I understand that it is used at the root zone level, the anycast technology. I'm not very sure about whether it is being used at the top-level domain level as well, TLD.

ALAIN DURAND: You can apply it anywhere you like, essentially. There are some limitations to the technology. It works really, really well for anything that is session-oriented and short-lived. You make one request; it comes back. If you ask the root, get some information off of the root, it's a relatively small packet. Usually it fits in one packet. It can fit in more, but it's a very small exchange.

If you ask a resolver, that's the same thing. You ask a question: "What is the IP address of that name?" There's a response that comes back, short response, and that's it. It's not a constant communication. Anytime you have such a pattern you can use it.

If you have a communication pattern but you have multiple exchanges, that's not as good because if I start again a conversation with Steve here and with David, and I say, "You are ICANN or you are somebody from ICANN. What do you think about XYZ policy?" and you start to tell me XYZ policy, but all of a sudden the communication drops and my packets get redirected to David. David may have a different opinion about what it means, or he may have a different perspective. So the communication may be a little bit surprising if I start with someone and complete with someone else. In that case, anycast technology is not necessarily the right choice.

UNIDENTIFIED MALE: So it's much more suited to the UDP rather than the TCP. Is that what you're saying?

ALAIN DURAND: No, it works with TCP, but for short-lived TCP connections. If it's a long-lived connections, that's not really a good idea.

UNIDENTIFIED MALE: Any other examples, apart from DNS, in which anycast is used?

ALAIN DURAN: David?

DAVID CONRAD: In the context of DNS service, since DNS primarily relies on UDP, almost any use of DNS can make use anycast. I know people use it for resolvers. I know people who use it for their authoritative servers. All it takes is the ability to build out or make use of the infrastructure.

There are other entities that use anycast. CloudFlare that I mentioned previously does most of their content delivery network services via an anycast cloud, and that's almost entirely over TCP. And, again, it just sort of depends on the duration of the connection.

The reason that becomes an issue is if you're having a communication across the Internet and the routing system changes — if there's a wire cut somewhere that causes routing to shift — then it's possible that the connection will go to a different endpoint, and since that endpoint doesn't know about the communication, it will just reset and the connection will be closed.

The only DNS service that I can think of that generally does not use anycast are zone transfers, because the zone transfers tend to be long-lived, a lot of data transfer. It would probably work, but there's no real need to do that because zone transfers don't happen that frequently.

UNIDENTIFIED MALE: Alain, I have a question from online. The question from Michael Oghia. He says, "If cooperation is so stressed, can you address what is hindering further development and deployment of IXPs? Is it mainly pushback for competitive/economic reasons?

ALAIN DURAND: No. I have seen a lot of IXPs being deployed in the last many years. I'm not an IXP person, so I cannot really talk about the challenges. What I know is that usually you can set up some very small IXPs for very little money, so this is not necessarily the main problem. In some places there are some regulation issues that could get in the way, but technically it's relatively easy to do.

David?

DAVID CONRAD: My experience — I'm not really an IXP person as well — but in my past experience, the largest constraints on deploying IXPs were either a regulatory environment within the country that disallowed the services that an IXP provides or in some cases just the environment in which the IXP is being deployed didn't really support a large [network].

If you have two ISPs in a country, having an IXP may not make a whole lot of sense because that turns into a single interconnect between the two ISPs. IXPs matter most when you have a large number of ISPs in the country that are trying to communicate to each other.

If the IXP does not exist, then you have to set up these bilateral communications with every ISP and you can rapidly use up your port and routers just by having interconnects with other ISPs, and particularly in cases where ISPs are not interconnected and having to use external bandwidth to communicate.

There were cases a long time ago in which I know of a country where the regulatory environment disallowed a direct interconnect between two universities that were in the same city, and they were actually traversing half the planet to talk to each other. They would go from their country all the way to the east coast of the US and then back to their country in order to exchange traffic. At this point, this was back in the early '90s. This was insanely expensive bandwidth and was just crazy, and eventually the universities were allowed to interconnect, but initially they had to traverse the entire globe to talk to each other.

ALAIN DURAND: Thank you.

UNIDENTIFIED MALE: Thank you.

ALAIN DURAND: Any further questions? Going once. Going twice. Lunch time. Thank you very much.

DAVID CONRAD: So once again, we're soliciting feedback. I know that many of you were in the last session. We're asking for feedback for each session. If you would please take a moment and go to this URL and provide some feedback and let us know how the session went, how it could be better, if you have any other ideas on technical sessions that we could be doing in the future.

We will be coming back at 2:00 pm for Ed Lewis's talk on protocols that top-level domains registries use. We look forward to seeing you guys back here at 2:00 pm. Please enjoy your lunch.

[END OF TRANSCRIPTION]