HYDERABAD – How It Works: Internet Networking
Friday, November 04, 2016 – 11:00 to 12:30 IST
ICANN57 | Hyderabad, India

STEVE: …into depth, because I will get any of his accomplishments wrong, and he's got many of them. So I'm going to pass it on to Allan. A little housekeeping, same format as before. If you have a question, please raise your hand, I'll come and bring the microphone to you, and Allan can address the question as you have it. We try to stay very informal here, so you know, please feel free to raise your hand and ask a question. With that, Allan, welcome.

ALLAN: Thank you Steve. So my first accomplishment of the day is to find the mute button. How many of you are lawyers? None? Okay. Well, initially I wanted to have a session that was IP versus IPR, internet protocol versus intellectual property right. So, when the community was sometimes, we only look at the intellectual property rights, without necessarily digging too much into the details of our technology underneath work.

And this session is all about doing this deeper dive into the foundation of the technology. So, we're going to talk about

naming, addressing, and routing. So next slide please. Thank you. And, first off, as an introduction, we are going to talk about the different layers of networking, and we will use the OSI seven layer model to guide us. This is not something very rigid, but something that helps us drive a discussion.

And as we will see, the seven layer model evolve into nine layers, and we talk about each of them. So, next slide please.

Next.

And feel free to interrupt me and ask any question at any time. So, when you talk about the layer zero, out of seven layers, from one to seven, we really talk about the physical [inaudible]. And the decision that one make when you are building a network, at that very layer, have a lot impact on how things are going to be working later. So, you have two phenomenal type of technologies. Have a wired technology, or a wireless technologies, and we have different characteristic. So, wired means, you have a copper, or fiber, and wireless, you have antennas.

So, when you have copper or fiber, to lay this out, you need to dig a trench, and that can be fairly expensive. So depending on neighborhood, there is more or less expensive depending on regulation, you have to do more or less paperwork. So, the cost of putting this in the ground can be fairly high.

Wireless is easy to deploy. You but an antenna and all of the sudden, you have an area that's being covered. But you still have problems. You have spectrum issues, so you need to get license of a spectrum. You also have environmental issues. Like nobody wants to have an antenna in his neighborhood.

So an upscale neighborhood or an historical district in cities, nobody wants to see an antenna on top of a very nice building like this. So, this is a bit of an issue. The wireless networks can address, very quickly, a lot of people and can serve a lot of people. But the bandwidth for a person is usually relatively limited.

On our wire networks, we can get a lot more bandwidth, but it goes to only a certain point. So, we still can be complimentary, for example, do you have an antenna that is served by a fiber so that you get lots of bandwidth to the antenna? And then you have maybe [inaudible] to have this bandwidth distributed among subscribers.

So for the rest of the talk in the layer, one to seven, I will mostly focus on wired, but most of it can be also translated to wireless. Next slide please.

So, now let's say we have a fiber, and we want to communicate over that fiber. So, fiber is a point to point link, meaning you leave the fiber at one place, and you receive light on the other

one. For some kind of [inaudible], like [inaudible] that we see the light and we transfer it to electricity, like.

So, at the beginning, we had white lights and fibers, and there was one possible communication for one fibers. You want two communications, you need two fibers. Not necessarily the best. So, in order to make things better, what we have started doing is to, instead of sending white light, send colored light at a specific wave length.

So, get you some color, so [inaudible], what is your favorite color?

UNKNOWN SPEAKER: My favorite color is blue.

ALLAN: Thank you. And let me ask someone in the audience. What is your favorite color?

UNKNOWN SPEAKER: Orange.

ALLAN: Orange, perfect. Blue and orange are very different. So, if you have a network, you have some communication you want to

send through this network, we will send it on the blue wavelength. For years, my customer, [inaudible] orange wavelengths.

If you had said blue, I would have a problem. So yeah, maybe not blue. Let's try green or orange. So, that way we can multiply a number of communication over the same fiber. And that's really, really interesting. If we want to translate this into radio, it's simply a different frequency.

So, we can have wireless on different frequencies, so you have different customers. Next slide. Now, let's say that I want to have a fiber between here, to have some communication between here and Hyderabad and my office in Washington, D.C. I'm not going to dig a trench all the way from here to Washington, D.C, every time somebody wants to have a new communication. That makes absolutely no sense.

So, we are going to reuse fibers that exist somewhere. And the way to do that is to connect those fibers together with some kind of active electronic ware, and create what we call a fiber path. So maybe there is a fiber that exists between the [inaudible] and some kind of a central office. And I have a fiber that may go from there to another city in India.

And I have a fiber that connects to the shoreline, where we can [inaudible] undersea cable that will send me my traffic all the

way to maybe in Japan, and then through the ocean to the west coast of the U.S. And then I can [inaudible] some terrestrial fibers, and like that build a path all the way to Washington, D.C.

So by interconnecting all of those layer one fibers, and many have to change colors, so I can start with blue here, but maybe later on blue is already used, so I have to change to green, and then change back to red, and different types of colors. But if I map all of this, and I configure all of the different equipment to essentially create a path, I now have an optical path from here all the way to Washington, D.C.

So for me, as a customer, it looks like I have one fiber. I don't see that there are multiple fibers. So that's really interesting. Now, there is some downside to that, which is, let's say that I want to have another communication for redundancy between here and other ICANN office in Los Angeles. And I'm going to say, I have two different fiber paths. Yes, but then they use the same actual fiber.

They may use the same undersea cable, or they may use the same terrestrial cable. So if there is a cut somewhere on this cable, both of my path are damaged. And I may have layer three or layer two [inaudible], I may have, but underneath, because we use the same actual fiber, and I don't know about it, actually

I'm in a bad position. If there is a problem, I cannot deal with that.

That's one of the issue when we have all of this multiple layer of technologies, is that when we don't really know what is underneath, we may make decisions about redundancy that seems to be good decisions, but in actuality, they may not be that good, because we don't know how it's implemented underneath.

So, something to keep in mind. In terms of speed, a fiber can be quite fast. We started from [inaudible] …very common now, 25 [gigabyte?], 40 is more data center speed, and [inaudible] now. It's very common for service provider, we have multiples of 100 [giggy?] to go over with maybe a terabyte per second. It can go really, really fast with fibers now.

Interestingly enough, also with copper, we can go really, really fast. We can get copper of [giggy?], no problem. Copper 10 [giggy?], a short distance, it also work. So don't think that 10 [giggy] means fiber, we can do that also with copper. It all depends on the distance.

Next slide please. So when we build all of this network, the world is not flat. So we don't want to see, this is just one gigantic network. We want to separate the networks into what we call administrative domain. The old days, we called them

broadcast domain. So if something goes wrong in a domain, it doesn't impact what is happening on the other side.

And that's why you need to remember that internet spells with an uppercase I, not lowercase I. It means a network of network. So you do things in your network, I do things in my network, we [inaudible] and that's fine. We all participate in the global internet, but what you do in one location, should not have much impact on what you do in another location.

At least, that's the theory. So, that's why the IP is the Internet Protocol. Again, this is not intellectual property, okay? Next slide please. So, now I have a way to send packets. Do packets are data-grams, the 1500 byte long at a maximum. And I can send those packets from one place to another.

So if I want to send a file, what I have to do is to chop the file into some packets that are 1500 byte long, and send them all to my destination. So we're going to role play this with [inaudible]. So, the first thing that we are going to do, is to establish a communication. So slight twist from yesterday.

So, I want to talk to Katie. And I want to make sure that she hears me, and I want to make sure that I can hear her back when she reestablishes communication. So the first thing I'm going to do is to say hello. Hello, Katie.

UNKNOWN SPEAKER: "Ack."

ALLAN: What just happened here is, I sent my first message. She heard it, and send some kind of acknowledgement back to me. So by doing this, now I know that she can hear me. If she had not heard me, she would have not sent anything back, right? The fact that I received "ack," means she can hear me.

Now, from her perspective, she can hear me also, right? But she doesn't know if I can hear her. She send me a packet that, at this point, she doesn't know if I have received it or not. So I have to do one more thing. I have to send a reply back to Katie, said, "Yes, I've received your 'ack' acknowledgement." "Ack."

UNKNOWN SPEAKER: "Ack."

ALLAN: So the last one is actually superfluous. We don't need a fourth one, because I already know that she can hear me from the first one. So that's a three-way thing. I send the packet, she respond to me, so I know that she has received it, and I send it back to her, so she knows that I have received her answer.

That's what we call the TCP, three-way handshake. So now that we have this communication established, what we want to do is to send packet. So, I have my big file again, and I'm going to send it packet by packet.

So, [inaudible] here is packet number one.

UNKNOWN SPEAKER: "Ack."

ALLAN: Now, I know she has received it. I can send packet number two.

UNKNOWN SPEAKER: "Ack."

ALLAN: Good. Send packet number three.

Silence. She has not received it. I don't know why, maybe there was some kind of cosmic wave that destroyed the packet. Maybe there was congestion in the network, [inaudible] somewhere had to drop traffic. So the packet was dropped and she never received it. Or, maybe that she received it, and maybe "ack," but the acknowledgement never made it back to me.

In any case, from my perspective, she has not received the packet. So what I'm going to do is, we transmit it again. Packet number three again.

UNKNOWN SPEAKER: "Ack."

ALLAN: Now I know she has received it, and I can go on like this. So, when we do that, at the same time we will adjust for speed, so that I don't send stuff too fast, make sure that she can receive it. So that's really what TCP is all about. There is another protocol, [inaudible] called UDP, but send data, but doesn't have this control flow.

It's only I sent, I receive, I sent, I receive. So there are cases where it's really, really important to know that, [inaudible] has received my data. For example, if I sent, I don't know, a picture, I want to make sure that she has received the picture completely. And it's not just half of a picture with my head chopped off, right? Now, if we have a voice communication, sometimes it's better to have a bit of a drop. Maybe I don't hear a word or two, rather than to have all of the sudden some pause, and some delay, and then I have a part of a sentence that was pronounced

maybe 30 seconds ago, because that would drive me completely nuts.

So this control flow may not be needed. In that case, what I simply want to do is send the traffic, and I will use the UDD protocol. So, we design a protocol depending on the type of communication and the intent in terms of reliability. Do you want to have absolutely reliability of a transfer, or do you want to make sure it's more real time? Then if you want real time, it's UDP.

If you want control of integrity of a transfer, you will use TCP. There are other transport protocol [inaudible] IETF, but those are really the two that are the most widely used now. Next slide please. So on top of that, especially if you use UDP to have real time data, you need to have some kind of a real time streaming protocol, that will describe, okay this is the type of data that we are going to send.

This is the bandwidth that we want to send it to, the situation of overflow and all of that. So, this is a description of this, that with RTSP protocol does. So, moving on to the layers of OSI model, from [inaudible] size from one to seven, we started at zero, we're now at six. Next slide please.

So, when you send data… If I send like binary [inaudible] data, I need to make sure that you understand what it means, because

it [inaudible]. So she needs to be able to pass this data. So, sometimes it's just a private agreement that we have. Sometimes we need to actually describe it more, to say this is a format of what is in the packet.

So years ago, packets were described using AS and one, which is a binary format, and it's very difficult to read. Later on, we use XML, which is a text format, like a mark-up language, like HTML, very similar. And now what is more fashionable is to use the [Jason] format, that looks like an example here on the screen.

It's easier to read, it's more like a [inaudible] type syntax, a dictionary type of syntax. It's higher level semantic. A lot of people like this new type of syntax now. So that will make sure that when I send something, I can format it in a certain way, I can describe my objects, for example, and send it to [inaudible], and [inaudible] will receive it and know exactly what this is about.

There is no doubt. So, all of this, essentially, are technical layers that enables something. Next slide please. It enables this. It's a kid on his bed, goofing around, and watching some videos on his iPad. None of the technology below matters to the customers. The only thing that matters to them is this. Can I watch my video?

The role of technology, is to make sure that this works. This is the ultimate customer. So, to do that now, most protocols are based on the web, HTTP, that's the first version of the web. HTTPS, second version of it. But really make the kid happy should be our goal.

So those are the seven layers of the OSI model. So we can go to the next slide to see the eighth layer, which is none of this matter. Even making the kid happy. None of this matter if it doesn't make financial sense. No service provider will deploy any infrastructure unless it knows it can make money.

Nobody that is deploying services over the top, will deploy anything unless they know they can make some kind of money. So, when you are making a business plan, if at the end of the day, this is negative, you will not do it. You will never get the financing to do it. So, that's a really important element to keep in mind.

And of course, because we are at ICANN today, next slide please, there is a ninth layer, and this is where we are now. This is a political layer, where we need to understand, how those things get organized. And let's have a policy discussions that we are having this weekend and early next week, those are really, really important to make sure that people get along and agree on what needs to be done.

And political doesn't mean that you have fighting, and bad elections, or crazy elections, like we have now in some countries where I live, but this is about getting along and finding some agreements on how we are going to be able to [inaudible] society. And that's the same thing for this community. That's why we have meetings at ICANN.

So those are the, there are zero from nine on this seven layer [IOS] model that I wanted to explain today. So those are the foundations of what we are going to do here. The next slide please.

Now let's talk about naming, addressing, and routing. So this all started when I woke up in the morning, and I had a really raging toothache. It was really, really, really bad, and I needed to do something about it. So, I went to see [inaudible]. Have really, really bad toothache. I need to do something. Can you help me?

UNKNOWN SPEAKER: Yes.

ALLAN: I need the name of a dentist here. Do you have a name of a Dentist who will work on me?

UNKNOWN SPEAKER: My dentist's name is Dr. Walker, Johnny Walker.

ALLAN: Good. And maybe this Johnny Walker guy can help me and I have a bit of his prescription. So let's try to find this Dr. Walker. Next slide please. Okay. So what is a name? I ask you what is the name of your dentist? So if you look at the dictionary, the name is a word, a set of words by which a person, an animal, a place, or a thing is known, addressed, or referred to.

Example, my name is Walker, Johnny Walker. So, if I know the name of somebody, if I know your name, I know who you are. First thing I wanted to remember, from all of this [inaudible], if you forget all of my layer zero to nine things, one thing to remember is that. What is a name? if I know your name, I know who you are. Next slide please. So, when I have a name, I can do two things about it. We can talk to someone, like earlier I was talking to [inaudible], or we can talk about someone.

So she told me about Dr. Johnny Walker. So I can go to Steve, Steve, I've heard that this Dr. Johnny Walker here is a dentists and is a good one. Do you know him?

STEVE: Oh yeah. I really know Johnny Walker well.

And his brother, Jimmy Bean.

ALLAN: Yeah. So, what just happened here is, we were talking about Dr. Johnny Walker, right? We're not talking to Dr. Johnny Walker now, he's talking about. So there are two functions of a name. Talking to somebody, or talking about somebody, having this conversation. So, when you pass the name from one party to another, we call this a referral. So [inaudible] gave me a referral for Dr. Johnny Walker.

So, next slide please. Names are scopes. They could be ambiguity. Example is, my name is Allan. I'm the only Allan in my family. So when we have a family gathering, somebody says, where is Allan? I know it's me. But Allan is a very, or was a very common first name when I grew up.

So when I was in school, elementary school, there were four of us whose name was Allan. So sometimes, the teacher would come and say, Allan, do this. And the four of us were looking at her and says, which one? And then the teacher would get really angry because nobody will do anything, right?

And then she would have to say Allan [inaudible], do this. Okay, now I know it's me, I can do that. So you see, you need sometimes to qualify the name, to have something that is

unique. If it's not unique, there is confusion, you don't know what to do. Next slide please.

But a name is not enough to communicate and do something. Now, I know about this Dr. Johnny Walker. I have no idea where he is. How to find him. So I need to go back to [inaudible]. I looked around, and [inaudible] for Johnny Walker seem to be a good guy. Now, [inaudible] and I really need to go see him. Do you have his address?

UNKNOWN SPEAKER: His address is 125 Root Canal Drive in D.C.

ALLAN: Thank you. So, now what did [inaudible] do? She went through her address book, and look it up, look up Johnny Walker, and found the address. She doesn't use her Rolodex anymore, she has a computer to do that, because she's a very modern person, but that's essentially the same thing.

We have cards with a number of entries. So when we talk about the DNS, in the previous session with Steve was explaining how to manage his own domain name, this exactly what this is about. We get the name, and from the name, we get an address. We can get other things. For example, it could be a telephone

number. It could be some notes about, is he a nice guy? Is he not a nice guy? Does he like X, Y, Z?

So, same thing in the DNS. We have different record types to describe all of this. So this process is called name resolution. We go from a name, and we go for the database, and some more data is returned to you. Next slide please.

So in the DNS space, there have been a number of discussions. This is an old technology, but more recent discussion about international organization. As I mentioned, I'm French. I cannot use the ASCII code very well, because we have different characters in French, like C with a [inaudible], E with some accents, and all of the letters like this.

Like I with a double dot. So, I'm sure any Indian, this is the same thing. You have different characters that are not ASCII. In China, you have Chinese characters. In Arabic, you have different character set that you write from the right to the left instead of left to right. So, if we want the internet to be a worldwide project, we need to make sure that we can handle those internationalized names.

Another recent activities… Recent, by what I mean the last 10, 15 years, has been to do DNS authentication. So the technology in SSEC, as pointed out earlier, this is really more about authentication rather than security. And maybe the problems in

the path when you send the, when you do the resolution of a DNS day, but at the end of the day, when you receive the data, you can validate and authenticate that the data is really what it should be.

And if it's not, then you can either dump it, or do whatever you want with it. For the activities was the expansion of the root zone. Initially, there was only country codes, and then dot com, dot net, dot EDU, dot org, but now you have hundreds of other names to choose from.

So, the more recent activity that have taken place in different forums like IETF or ICANN. Next slide please. Okay, now I have this address, 125 Canal Rd. This is the address of the dentist. Next side.

So, what exactly is an address? Again, looking up in the dictionary, those are the particulars of a place where someone lives, or an organization is situated. Remember, if I know your name, I know who you are. First thing to remember. Second thing to remember is, if I know your address, I know where you are.

Those are two different things. Who you are, where you are. So an address is where you are. Next slide please. So let's look at some example of addresses before going into the IP addresses,

because they are essentially [inaudible] same searches. So let's look at, for example, as I said, I live in Washington, D.C.

And there is a very famous small, white building here that everybody wants to live in, and this address is 1600 Pennsylvania Ave. NW, Washington, DC 20500-003, USA. It's a very nice building actually. If you look at the address, there is a hierarchal structure. You have to look at it from the right, and move back to the left.

USA is the country, 20500-003 is the zip code. DC is District of Columbia. It's not a state, it's like a difference in the US. Washington is the city. Northwest is the quadrant. DC is split into four quadrants, northwest, southwest, northeast, southeast. And Pennsylvania Avenue is the street. 1600 is the house number.

Not all addresses are structured in such a geographical way. For example, again in the United States, you can dial a toll-free number, meaning you dial the number and you don't pay for the communication. The person on the other side pays for it.

We start by 1-800. And when I dial a 1-800 number, I have no idea [inaudible], I have no idea where the person is. It could be somebody local in Washington, it could be somebody in New York, or Chicago, or Dallas, or San Francisco, or the call can even

be forwarded in India, or in China, or somewhere else. I don't know.

By simply looking at phone number, I have no idea where the designation is. Cell phone numbers are exactly the same in the US. You get your cell phone number from where you live. For example, I live just outside of Washington, and the area code is 703, but if I am here in India, you can still call me on the same number.

If somebody calls me from the US, thinks maybe I live in DC because I have a 703 number, but I happen to be in India. So by looking at the phone number, you have actually no idea where the person is. The same thing with the IP addresses. By looking at an IP address, you do not know where that IP address is located. There are some other database, that people have build manually, those are called zero location database, where you can look up an address, and then we say, this is somebody in this particular city.

Sometimes it go down all the way to the streets where you live. But this is not by simply looking at the structure of the address. This is something that we have build, by using all kinds of other techniques. So next slide, please. So, as I was mentioning about names that have scopes, same thing, addresses have scope.

If I am in DC and I ask somebody, where is the White House? What's the address of the White House? They would say 1600 Pennsylvania Ave. NW, they will not mention DC, right? Because it's assumed that we are in the same city, we know that. But sometimes, it's confusing.

If you are in Europe and I ask you where is Paris? You will say it's in France. Probably if I ask somebody here, where is Paris? They will say it's in France. But in the US, there are 29 cities called Paris. There is one of them that is not far from where I live, [inaudible] Drive, it's very small village, maybe, I don't know, 50 people live in there.

So sometimes I joke with my kids, we are going to Paris this afternoon. Let's just drive there. And they understand that it doesn't really work like that. But, again, with an address, it can be a local address, but if you want this to be global, you have to qualify it. Next slide please.

So we can use an address directly, for example, we can send a postcard, or we can just like names, pass the address along to somebody else, as a reference, to talk about it. So [inaudible] told me the address was 125 Root Canal Road, that was a reference to this address. I come back now to Steve.

Steve, I'm going to see Mr. Johnny Walker tonight, and it's on Root Canal Road, and I'm not quite sure. Is it a good district? I mean, is it safe to go there?

STEVE: It's kind of safe. You might need a little bit of security.

ALLAN: All right. So, what just happened, we had a conversation about this address. And it's the same thing with IP addresses. We can have a conversation about IP addresses. For example, there are IP addresses that are well known for sending spam. And people will say, we are going to filter out those addresses. So we have a list of those IP addresses, put that into some kind of a black list. We have this conversation about this list. Okay?

So the IP address is not only used to communicate, but it's also used to have that conversation. Now, I need to go to the dentist. I just can't hold it anymore. I really need some of his prescription. So, let's go to 125 Root Canal Road. Next slide please.

But how do I get there? Just keep on with my parallel with sending a postcard. Let's say that I want to send a postcard to the White House. What I have to do is to write the address, 1600 Pennsylvania Ave. NW, Washington, D.C., etc. And I can post it

here in Hyderabad, and then a few days later it will show up in Washington.

How does it work? Is it a miracle? No. It works because there is a post system in India that's going to pick up the mail in the mailbox, and parse it, and say, oh, this is something that's for the USA, and maybe they will send this through an airplane to New York, or to Los Angeles, and at some point, it will be handed over to the US post system that will send it to Washington, and then distribute it to the right place.

This works because there is a collaboration between these post systems. There was no collaboration, it will never work. So for the internet, this is the same story. We need this collaboration to make sure that we can go somewhere. So how do I go to 125 Root Canal Road? Next slide please.

Before going there, we have been talking about postcards, let's bring the discussion back to internet addresses. There are two protocols that exist on the internet today. You can say two is a lot, but in the past we used to have much more than that. We used to have things like Apple Talk, for example. We had [inaudible] network, we have [inaudible], we had IDN [inaudible] network.

So there were lots of different protocols in the past. Over the last 10, 15 years, we went back to only IP. Now we have only two

versions of IP, we're diverging again a little bit. The version of IP, whichever we're using, is called IP version four, it was defined in 1981. The newer one, it was defined in 1994, [inaudible] in 98 is, IP version six.

So the question I often get is, what happened to IP version five? Well, when the standard was developed and agreed on at IETF, somebody went, it was [inaudible], actually he did in fact, he went and he looked into the registry that counting the IP version number, and he realized that version five had already been allocated for [inaudible] protocol.

That was about multicast. And sending real time data over the network. So we could not use version five. The next one available was version six, that's why it's IP version six. Nobody knows what happened to version one, two, and three, by the way. Lost in time.

So if you look at those two protocols, they are mostly the same accept for the length of the addresses. In IPv4, an address is 32 bit. In IPv6, it's 128 bit. So this is not four times more addresses. It's a lot more than that. It's actually square and square. So we start from three billion addresses, and you end up with this gigantic number here.

The goal was to make sure that for the next 50 years, we would not need to go to yet another version of IP. Will we be successful

or not? We'll see. But from the way the protocol works, there is very, very little difference between the two, except that, because the addresses are not the format, they cannot communicate with each other. And we see, it has been a bit of a problem.

Next slide. So, I'm sure you've all have heard about IPv4 exhaustion. Well, this is not that the address gets tired. When you have an address, it works today, it was working yesterday, it will work tomorrow, right? When we talk about exhaustion, actually, what it means is that you can't get any new one.

You used to go to the IRR, to the version of the virtual Internet Registries, and says, well, this is my new network, my new plan. Please give me a number of addresses, and then you have a discussion about how many addresses you really need. And this was coming from a pool that was called the [free?] pool, that was handed over to the IRR by the IANA.

This pool, this [free] pool, is gone. So each of IRR, have enacted policy on what to do with whatever addresses left they have, and each of them have a slight variation of this. But the internet is going, so what's next?

Well, the first answer to the what next was IPv6. That's the new thing, just use the new thing. Well, didn't really work that way, because with two are not compatible. So if you build a new network with just IPv6 and nothing else, while it's great inside of

your network, you have numbers, you can use it, but when you want to communicate with somebody outside that has the old system, it doesn't work.

So, that's a bit of a challenge. Next slide please. They are not compatible. It's almost like trying to put a US plug into an Australian socket, it just doesn't work. We can put adaptors, and with a MAC address network [inaudible] from IPv4 to IPv6, but when you do that, we need to put addresses in the pool, and if they are translated from v4 to v4, or v4 to v6, we use the same number of addresses.

So having v6 doesn't really help the equation here. Next slide please. So, for a service provider to deploy IPv6 only, and nothing else, no translation mechanism there, nothing else, is a risky proposition, because your customer may have devices that won't support it, and then won't use it, or they may want to go to this nation that don't support it, and that's also useless.

So, a colleague of mine was travelling last summer, and he went to a hotel in Germany, and that hotel had an IPv6 only service. I was really surprised when he told me that story. Wow, somebody who is quite adventurous. And how does it work? So I ask him the next day, so how did it go? He said, well he had a laptop like a Mac, just like mine, and most laptops have support for IPv6, so it was fine. Could connect to a network, he could go

to Facebook, he could go to Google, but he could not go to the news website that he used to go to.

He could not go to the website of the school his child was going to, and a bunch of other places. So his experience was, yeah, it's not exactly the internet, it's limited. So this is not exactly an interesting value proposition. In the meantime, IPv6 gets more and more deployed, and you have places like Belgium now, where it's about 50% of traffic, that's IPv6, it's quite remarkable.

It's not the same all over the world. In many countries, it is zero. I have made some studies where we look at exactly which countries are developed, and is there a correlation with the GDP per capital of those countries, and the answer is sometimes yes, sometimes no. But this is a work in progress, not quite there yet, not quite ready to be a direct replacement.

What this means is that you will have IPv4 and IPv6 coexisting for a long time, and by that I mean not simply a couple of years, maybe five years, maybe 10, maybe even more. Next slide please.

So in the meantime, you still want to go, your network, how do you do that? All five IRRs have enacted new policies to allow transfer of addresses. So what is a transfer? A transfer is, if party A, it has a number of addresses that it is not necessarily using, and party B who wants more addresses, they made some

kind of private financial contract, and then they go together to the IRI and says, please [inaudible] this block, transfer it to party B.

So what the IRR do is, check that party A is right, [inaudible] of the address block, and that party B somehow qualifies for this address space. And that's it. We're not aware of the final part of the transaction, so you don't have any data about that. So the details may vary, depending on which IRR you are.

You may also have transfers from one IR to another one. Most of the transfer concern what we call legacy addresses. These are our last blocks that were allocated prior to the existence of the IRR system, in the late 80s, early 90s, very last blocks. And most of those blocks, not all, but most of those blocks are in the North American region.

So now we see a last flow of addresses coming out from the ARIN region to other regions in the world. We see also some [inaudible] that is more interesting to observe. But now what is happening is, we have a global market, where addresses are flowing essentially from any country to any country, minus some policies that will prevent them in certain spaces.

So, this is the way people are building new network now. The cost of this is around $10 an IP address. And it went down for a while, now it's going back up. [Inaudible] depend mostly on the

size of a block that you're buying. Small blocks, the price for IP addresses is higher than larger blocks. It's normal like retail versus like [inaudible] wholesale.

On top of that, acknowledgements like NAT, network address translation, are a higher grade version of that. A service provider can enable to leverage the space. For example, some of you want to transfer and acquire a block of slash 16 in IPv4, which is 65,000 addresses, with 65,000 addresses, you can have 65,000 customers, that's not a lot if I want to start a new wireless service. But if you do multiplex with NAT, and have 100 customer per IP address, which is fairly reasonable to do today, now your 65 number translate into 6.5 million. With 6.5 million potential customers, you can have an interesting business plan.

So this is what is happening in the world today. Next slide. Skip over that, this is not the most interesting here, let's skip over this. Those are just some statistics, how many addresses have been transferred, there is quite a lot. Just let's, next slide. Next slide.

The point I have made about IPv4 is that, and I want to re-discuss this, is, this is not an IPv6 should be, no. It is that while IPv6 is being deployed, it's really important to keep in mind that you need to maintain some kind of IPv4 service, and we need to have discussions about what to do with this IPv4.

Now, back to my dentist, because I remember that, back to the tooth that is hurting, how do I go there? So, let's say I want to walk there, or drive there, I'm going into the streets, and there might be signs on the street that says, root canal road is second on the left. Go to the roundabout, third street, first exit on the roundabout.

Okay? Somebody has put those signs before I start driving, they just don't show up, [inaudible] when I'm driving. That's the same thing on the internet. Somebody has to build routes on the internet with some signs that give me directions, before I can send traffic. So let's see how it works. Next slide.

So, route or route depending on which version of English you're using, is a way, a course taken in getting from a starting point to a destination. So let me step back. Remember, first thing I really wanted to remember. If you have a name, you know who you are. You have an address, you know where you are.

If you have a route, you know how to get there. Those are three different things. Who you are, where you are, how to get there. If you remember only three things today from my tutorial, those are the three things to remember. Who you are, where you are, how to get there. Next slide.

So let's build those maps. So I have drawn here network topology with a bunch of links, and a bunch of routers. I'm the

source on the left, I want to send traffic on the destination of the right. So as I explain earlier, first before I send traffic, somebody has to build those routes. How is it done? Well, it's done from the destination back to the source in reverse.

So, the service provider that is connecting the destination knows how to get there. He has a joint link to the destination. So he's going to advertise, to announce, to all his peers, I know how to go to Root Canal Road.

The peers receive this information, and are going to send this to their own peers saying, I know somebody who knows how to go to Root Canal Road. And so on. So, routing is all about, I know a guy who knows how to get there. Or, I know a guy, who knows a guy, who knows a guy, how to get there. This is the same collaboration system that I was describing with the post office.

Without the of collaboration of the service provider to go and build those routes, there will be no internet. There is no central authority here. It's all about peer to peer collaboration between the service providers. So, when I'm going to send traffic, next slide please, once those routes have been established, I'm going to send this to my service provider that finds the routes to go to another hub, sending the packet there, and is going to expect that the next hop is going to do the right thing, which is forward the packet, to the following hop.

Same thing as in my post office example before. I send a postcard here, I expect that somebody is going to collect it and do [inaudible] send it to the main post office station, where somebody is going to do the right thing, and maybe send it on a plane to the U.S. I expect those right steps to happen because of that collaboration. Same thing here.

I expect my service provider to do the right thing and send the packet where it needs to go. And then the next service provider will do the same, and so on and so on. And again, this is [inaudible] part of system that is used not only to build the routes, but also to move the traffic around. This is really, really critical.

So, is this secure? The answer is, yeah, kind of, but not always. Next slide please. Sometimes there are bad actors. You can have a bad service provider, that has happened in the past. Or, you could simply have a bad guy who is going to inject himself in this conversation. And what he can do is to try to convince one of the routers, or one of the service providers in the middle, that he has a better path to the destination.

How to do that? Well, he is going to say his metric, his BGP metrics is better, he's going to make a louder announcement, or whatever it is, but if he convince the node that is the [inaudible] that he has a better router, than all of the traffic will be diverted

to the bad guy. So this has happened. Sometimes happened accidently, because there is a misconfiguration somewhere.

We had a very famous example of the traffic to YouTube in Pakistan a few years ago that was diverted somewhere, as it was just a misconfiguration. So, that can happen. I'm going to finish this slide, and then when we go to the question, if somebody has a question, to prevent from that, there is a system called [RPKI?], or Resource Private Key [Infrastructure?] that was developed, and the idea is a bit like DNSSEC, it's to use cryptography and certificates to validate resources.

So, when you get an address block, you can get a certificate that says, this address blocks belong to you. So when you send an announcement saying, this is my address block. You can reach me through there, you can sign it. And when they receive this announcement, or repeat of this announcement, can verify the signature of the announcement, and if it is [inaudible] when they can accept the announcement, if it's not [inaudible], then you can drop the announcement.

So in that case, when the bad guy was shouting, I can go to Root Canal Road faster, then you compute, the verification of the signature, and you realize that's the wrong signature. It doesn't work. Then you can drop it, and then you would prevent it from this type of attacks. There are still some issues with that.

The first issue was a political issue about, do we have a system that is centralized with a single source of authority for the certificates? Or do we have multiple source of authority for the certificate? And that's still a discussion that's ongoing. The second one, this provide validation of the origin, it doesn't provide the validation of the entire path. And for a number of service providers, that was simply enough. They wanted more.

So as a resource, [RPKI] is being used, but is not universally deployed. There was a hand, it was raised, there is a question online? Not yet, okay. So, next slide.

Well, now, I know who you are. I know that I need to talk, I need to go see Dr. Walker, Johnny Walker, right? I know where Dr. Johnny Walker is, 125 Root Canal Road, and I know how to get there because there is some signaling on the road that take me there. Those are the three fundamental things in communication on the internet: naming, who you are; addressing, where you are; and routing, how to get there.

So next slide. Then somebody can take care of my tooth ache. So now that I feel better, that I have a dose of this Johnny Walker prescription, I can take some questions.

UNKNOWN SPEAKER: Let me bring in the microphone. Hang on one second please.

UNKNOWN SPEAKER: Is there any security measures similar to RPK? Say, 125 Root Canal is down? Address of dentist. She can now, I also say that 126. Is there any, how can I [inaudible] that she's saying 125 or 126?

[CROSSTALK]

…how this RPK is so-called security measure? Is there any…? How can I [inaudible] that name my name server?

ALLAN: So, this is what DNSSEC can be of use. I can ask a question, and she can sign the answer, 125 Root Canal Road. And if somebody messes the transmission around, and what I receive is 126 Root Canal Road, and I'm going to compute the security validation of this, and I see that it doesn't match, I know that there is a problem, exactly the same thing as [RPKI].

Now, if it validates as 125, but she had the wrong address to begin with, nothing I can do about it.

UNKNOWN SPEAKER: Basically works on the trust name on the name server?

ALLAN: Well, if I ask you, what's your favorite color?

UNKNOWN SPEAKER: Orange, I said.

ALLAN: I trust that you're giving me, you're telling me the truth. Right? So, when you're asking your name server to tell you an answer, and you verify that nobody has messed up with the answer, you have to trust your server, because if you don't trust, there is no point in even asking the question.

UNKNOWN SPEAKER: [Inaudible]… Similar to the [inaudible] as you explained, some bad actor knows that this is the [inaudible] for the thing. Same thing can happen to name servers. So, this time I want to confirm, is there any mechanism to ensure that name server, whatever it is declaring, is it legitimate or somebody is doing?

ALLAN: You can verify, but nobody else tampered with data, but you don't know if the data was right in the first place, right? The same way, if I go to Root Canal Road, 125 and I go there, and somebody has, I don't know, kidnapped the dentist, and put a

fake dentist in there, I have a problem. That's the same type of parallel that you can make.

So [IPKI?] and DNSSEC could really be understood as the same thing, one for routing infrastructure, one for DNS infrastructure. Another question?

UNKNOWN SPEAKER: Yes. I am network engineer, but I have two questions. The first question, is that, if I understand…

ALLAN: Speak closer.

UNKNOWN SPEAKER: Yeah. If I understand you well, you said that IPv6 network can deal with IPv4 network directly, there should be maybe gateway or proxy, maybe something like that, that translate from IPv6 to IPv4 or opposite. My question is, how the larger packets of IPv6 will be translated into IPv4? And the opposite? That's the first question.

The second question, are the network elements are smart enough to detect that there is an attack coming from IP or range of IPs?

ALLAN: So let me try to answer your first question, I'm not sure I get your second one. So, when you do network address translation, it started with private network like, after 1980, network 10 dot [inaudible] 192 168 dot something, to global network. And what you do is on the box that does the NAT function, you remember the incoming packet, you create a mapping between the internal address and an external address, and you send the packet out.

So you can multiplex multiple communication from multiple sources, by simply playing with bot numbers. When you do a NAT between IPv6 and IPv4, this is exactly the same thing. Instead of mapping your private IPv4 address to a global IPv4 address, you map an IPv6 address to an IPv4 address. And the rest is exactly the same thing.

So, you need to put a pool of IP addresses on the NAT box to do that. And the size of the pool is exactly the same. Everything will operate exactly the same way. Now, your second question, will you please repeat it?

UNKNOWN SPEAKER: The first question, for the first question, [inaudible] before, [inaudible] net is uneven.

ALLAN: What?

UNKNOWN SPEAKER: Even. Not accepted. NAT is not accepted. [CROSSTALK]

ALLAN: …position that some people have, but I will make an observation that has been deployed, it's a technology that has been deployed and still being deployed, massively there are millions, billions actually, of NAT boxes being deployed. The last number of service provider that use that, and it works.

UNKNOWN SPEAKER: So, I am system admin, if I try to dig, to look up for a domain name using the date, maybe this domain name is in IPv6 server, I can [inaudible]. So I try to dig in the name servers that have IPv6 [inaudible], so I get to the [inaudible] from this server, about this domain name.

ALLAN: So that's a different problem. What you're talking about here is an issue in the DNS when you do the resolution that was described earlier by Steve, when you have all of those referrals, if you don't have the same version of IP on one of the DNS servers, you have a problem. So we wrote a document, a number of years ago, back in 2003, 2004, and I think this

document number is RFC 3901, I have to double check the number, it's a best practice document on how to handle this.

And what it says is, in order to maintain the capability in the DNS, when you have a DNS server, at least one of the DNS servers, you know, you can have many of them, at least one of them has to support IPv4. When you [inaudible] there must be at least one that has IPv4. If not, we know that there will be problem, and it is a misconfiguration.

Same thing with the resolvers. If you do an IPv6 only resolver, you may not be able to reach anybody, or very few people. Have a dual stack resolver. And that way you will avoid a lot of problems.

UNKNOWN SPEAKER: Thank you. The second question, my questions was, are the network limits are smart enough to take [inaudible] there is an attack like DDOS from an IP or several IPs? And the look at the traffic from this range [inaudible]?

ALLAN: So, that's a very interesting question. There are a number of solutions that exist, products, that do exactly that. So, they look at the traffic, they inspect the traffic, and they say, this looks like an attack, I'm going to drop this traffic. Okay? Now, it works if

you have up to a certain volume of traffic that can be handled by this box. But if you have 10 times more traffic than that, it doesn't work anymore.

For example, for root servers, that may or may not be the right solution, because sometimes it's easier to simply respond to the packet than to do this analysis if it's a good packet or not. Let's say that it takes 100 cycles of CPU to respond to the packet. And it takes 1,000 cycles to do the analysis. Is it a good packet or not? Maybe it's more cost effective to not think about is it a good packet or not, and simply respond.

So there can be some interesting subtleties on how you need to handle those DDOS attack. It's not always simple. Another question?

UNKNOWN SPEAKER: Here we go. Moving the mic over.

UNKNOWN SPEAKER: What efforts are being taken to move IPv4 to six, especially in hardware problems complaints? Globally, standard efforts because it is going through, it looks like.

ALLAN: This is not something that I can really, is involved into, because ICANN has a very limited mission. So we're not going to see other vendors and ask them to do something about it. I can speak to this, because I used to work for one of those vendors. And most of the big router vendors have, for at least 10 years, sometimes 15 years, have product that can handle IPv4 and IPv6. Some of them have better performance than IPv4 and IPv6.

Some of them have the same level of performance both. So, from a [inaudible] perspective on big routers, the problem is solved. On some other equipment like older [inaudible] for example, you still have a fairly large number of [inaudible] that are still in production.

And those have no great path to IPv6, and in that case, it's about replacing the equipment. Sometimes replacing by simple [inaudible], like moving to a newer technology, or going to the same technology, but simply faster, or product has been [retired] and it's time to buy a new one.

New equipment for [DDOS?] function exists in IPv6. When you're going to things like in the back office, that's where you have a lot of problems, because you may have back office solutions, for example, to go and provisions was [inaudible] of those modems, that may or may not support IPv6.

And I use those to work for a service provider, but [inaudible] exercise, to look at what needs to be changed. And it is a multiyear effort to do that. The older the infrastructure, the older the system to go and manage it are, the more difficult this is. The newer infrastructure, the higher your chances are to deploy this.

And I was mentioning and made a study that was covering IPv6 penetration and GDP per capital, and when I did that, we stumbled on some outliers. In a country like Bosnia, I was not expecting them to be high on the map on IPv6 deployment, but they are. And when we talked to some of the people there, they said, well you know, we had been through a lot of turmoil, and after the war, things had to be rebuilt.

So, it was a newer infrastructure, and because it was a new infrastructure, it was easy for them to do IPv6 as well as IPv4. So this is an interesting reversal of a situation that's happening there.

Another question? Maybe online?

PAUL WILSON: Allan, thank you very much. It's Paul Wilson from APNIC. I just had a comment, actually, about the issue of compatibility of IPv4 and IPv6, because I think there is a chance that a takeaway

from your account of the capability may be that there was a terrible misjudgment, and a terrible sort of incompetent effort on the part of the IETF to design IPv6 in a way that was not compatible with IPv4.

And I've heard it described that, as though some terrible mistake was made. And I guess there is a question about whether something different could have been done, but the fact that is, that at any one layer of the protocol stack, the protocols operating are often, and more often than not, incompatible, UDP and TCP are not compatible, Ethernet and [inaudible] 11 are not compatible on the wire.

This is a feature of the protocol stack, that at any one layer, you've got a choice of one thing or the other. And the compatibility issue of B4 and B6 could be described with an analogy of, for instance, electric power or gasoline power in cars. I mean obviously, electricity and gasoline are incompatible with each other, and cars that take one can't take the other.

But the cars actually still drive on the same roads, so there is a compatibility in terms of their ability to use the underlying infrastructure. They follow the same set of rules that make them interoperable, and they also provide the same services, say, to drivers and to passengers, and to those who use them.

So I just think, there was at the level of the protocol itself, sort of choices that could have been made, but it's certainly not that there was an accident or a terrible misjudgment, or a fault on the part of those who decided actively the v6 should not, does not need to be compatible with v4. Some things would have been easier had it been done that way, but there was an active choice, which I guess had to do with the idea that compatibility with v4 would have introduced some baggage to v6 that would have lasted forever, whereas a clean switch from v4 to v6 let the protocol be a lighter weight, more efficient protocol.

I wasn't there actually. You probably know more about the discussions that happened at the time, but there is probably a more nuanced in there, in terms of how the decisions were made, then people often appreciate. Thanks.

ALLAN: Thank you Paul. [Inaudible] a response to this. Because I was there for this part of the story. The thinking back then, was that the number of devices connected to the internet… Remember, it's 1993. Okay? Doubled like every 12 months. Why bother about the legacy? In 12 months we have, 24 months from now, the device that we have today will only be 25%. Who cares?

So when we were in this exponential growth, there was no real need to think much about the legacy, and we though that these

things would be deployed fairly quickly. All we had to do was to get into the major operating systems, like Microsoft Windows, [inaudible], IBM, X400, and a couple of others, Cisco, and we'll be done. Right?

It didn't work that way. It took much longer than we expected. And back in '96 was kind of a major stop, when we looked at the eight plus eight proposal, when we went from a slash 80 to a slash 64, and tried to reset that, and realized that it would take much longer to go and deploy, because already back then, it was still exponential but not with the same slope.

And come 2000, we've seen a lot of growth in the internet, but not the same growth. And now that we have billions and billions of devices, the argument that in two years' from now, this will be only a fraction of it, doesn't apply anymore. So that was true then, but because it took so long, it is not true anymore.

Now, I also wanted to talk about your analogy with cars, electric cars versus gasoline car. Yes, we can take the same road. However, if you don't have gas station, you cannot fill up your car, or if you don't have electric station, you cannot recharge your battery. So you need to have some supporting infrastructure in order to get your car along the same roads, right?

And that's probably the message I'm trying to give here, is we are at the point today where we cannot simply build an infrastructure to support electric cars, because we still have a lot of gasoline cars. We still need those gas stations. So let's not forget about those gasoline gas stations now. Maybe later we could, but right now, we can't. We still need to maintain those gasoline stations.

That's probably the right analogy here.

UNKNOWN SPEAKER: I do want to thank you because I've never considered my Prius to be dual stacked before, and I want to appreciate, because you're going to make my wife give me stranger looks than she already gives me now, so thanks for that, Paul.

Are there any other questions or comments for Allan? We have nothing online still. I think all of this talk about dentists that really they're just chomping to get at the desert tray for lunch. So, they're trying to withhold their questions so they can get first in line.

ALLAN: Well, thank you all very much for attending this workshop, and don't forget that we have more coming this afternoon.

UNKNOWN SPEAKER: Yes, our next section will be at 1:45. We are taking a break for lunch. At 1:45 we're going to have Francisco Arias from ICANN and Joe Waldren from VeriSign presenting topics on RDAP and EPP, regarding the registry operations workshop as well. So that's at 1:45. And then at 3:15, we have members of the root server community to come and present about some [inaudible] stuff, some RSAC root server stability advisory committee. And also talk a little bit about any casting, which is a different type of routing.

I think you might have touched base, I'm not sure if he did or not. It's an interesting concept, it works really well for DNS, and it would be interesting to here. And it's a good presentation. So please, enjoy your lunch. Thank you Allan for this session. And we hope to see you back at 1:45. Thanks.

[END OF TRANSCRIPTION]