India Telecom News

http://indiatelecomnews.com/hallmarks-of-the-impending-internet-revolution-us-ignite/

Hallmarks of the Impending Internet Revolution: US Ignite

June 1, 2013

Technology is a wonderful and interesting thing. About every 30 to 40 years, something major changes. It's an S-curve. If you look up, Everett Rogers has shown that the adoption of technology is at first slow, then very rapid. And then, at the conclusion, it tails off and another technology comes and takes its place. You can see this happening in television. First, black and white TV; about 40 years later, colour TV; about 40 years later, high-definition TV. Hopefully, it won't take 40 years to get to ultra-high-definition TV, but these things change every period of time.

And the Internet and Ethernet is about how old? So it's about time. It's about time for some major inflections and some major change, and I think you're going to hear all morning about some of those major changes, some of the folks who are leading those changes at some of the leading companies and other places where that change is happening. I think that the changes are going to really be profound in the sense that they're going to impact people's lives. And I have begun working with a non-profit organisation, US Ignite, on how these changes might impact people's lives.

When US Ignite was launched, and US Ignite is about new applications of these new technologies, we had John Holdren, the Head of the Office of Science and Technology Policy for the United States. We had Subra Suresh, the Director of the National Science Foundation.

We had Julius Genachowski, the Chairman of the Federal Communications Commission. We had Larry Strickland, the Head of the National Telecommunications and Information Administration. All of these folks were there to celebrate the launch of a new non-profit organisation. It's not a government organisation. It's a non-profit [stickening]. It's working with the government but works independently and works together with other parties, to go and launch US Ignite.

So what is US Ignite? As I said, it's a public/private partnership. It's government working together with the private sector in a non-profit way. And the major numbers up there, 501(c)(3), mean something to Americans who recognise that as the section of the Internal Revenue Service code that says that is a non-profit organisation.We have a major, simple goal, to take these new technologies and turn them into new applications to benefit citizens of the world. We want to see next generation applications. And you see there, in orange, the areas that are being impacted. At the top, clean energy. Going around clockwise, advanced manufacturing, public safety, transportation, healthcare, education, important things that impact people everywhere.

What can these technologies do for them? And we're working together with a number of communities. We work with industry, with government, research and education, foundations, communities, and other organisations who are allied in this goal to help make the next generation of technology turn into beneficial applications for citizens of the world.

We have three goals at US Ignite. The first is to create 60 compelling, transformative applications based on things you couldn't do today, based on new technologies, things that you've been hearing about and will hear about today, software defined networking, local cloud computing, taking gigabit to the end-user, reducing latency, things that are going to change the way the internet works today. Second, that we get 200

community test beds, 200 communities, who are eager and willing to adopt these new technologies and their applications. And third, to coordinate best practices among these communities, among the industry partners, and to make sure that government, industry and communities are working together to make this goal happen.

These are the partners for US Ignite. You'll see here some industry companies. You'll see here some communities. You'll see here a number of folks who are working together to make US Ignite happen. The next generation infrastructure I think will have three big, important spokes.

Number one is the use of virtualisation and software defined networking. Those are not quite the same, but they're close enough that I'm going to put them into one category for this morning. And for all those who will come later and say they're not the same, I agree, they're not quite the same, but I still put them in about the same category. These are things that are going to change the business model for how networks work and are charged for.

The second is local cloud computing, and I will speak about this in my talk. And the third is that we don't need necessarily more speed. We need more responsive ability to reach the end-user, and I'll talk about that. So I'm going to mainly speak about numbers two and three, because I believe other folks today will talk about that first one up there.

Okay. So I thought I would tell you my punch line first. The whole talk revolves around these three changes, first that today we think that more bandwidth is bigger, and I think that's going to change. The new version is going to be more responsive is better. And I use the word locavore. That word is used in the food industry to mean eating foods grown near you, and I'm using it to mean consuming computer cycles and networking located near you.

Second, today we think about putting things into a massive data centre, into a massive cloud, concentrating many, many, many computers all in the same place. I think that is going to change, and we're going to see the cloud coming to you to be more responsive to the applications.

The third is that today we think that if we have a big enough internet pipe, that will be fine and traffic won't conflict, and that's true. But I think over time, and especially for wireless, it's going to be much more important to use our resources more efficiently, and tomorrow multiple, dynamic virtual pipes will be purpose built, will be bespoke per application.

So those are the points I'm going to make. And now that you know all about them, I'd like to go and start with more bandwidth is better. I have an internet connection in my home which is about 20 to 25 megabits per second, from Comcast, and of course I would like it to be faster. But would it really do much good? I did an experiment.

Here's the experiment that I did. I went and measured how many web page loads per second I could get. On the right-hand side is my connection unaltered, and you can see I did this experiment a number of times. It's not always the same. You can see the number of web page loads there. Okay, pretty decent. It's about one per second, a little bit less than one per second. That's very typical. Then I went and I found, as I said yesterday, an old Ethernet hub in my closet which was only 10 megabits. And so I went and I put that into the circuit. I did a speedtest.net. It said it was about 6.5 megabits.

So I then took the same readings running it through that hub to slow things down, and you can see that I got fewer page loads per second, as you would expect. But I didn't get half. From speedtest.net, the speed tripled, between 6 and 22, but I didn't get three times the number of page loads, did I? In fact, it only went up 23%. So clearly something else is going on.

Why is it that when I tripled the speed I didn't get triple the page loads? Well, there's actually a very good reason. This is what happens between my computer and the place giving me all those page loads. There's all these routers and links in between, and in fact the only one whose speed I changed was actually the second to the last one, from my computer out. Now, my computer to my home router was the same fast

speed, but when I changed the link going to Comcast, my provider, that changed. So, only one of those links changed speeds.

Now, of course the weakest link will determine the speed, but as these things go, it had to go through all these other links in order to get to the server which was in a massive date centre. Where was that server? I ran a technical program which allows me to see where all of those things were. So I tracked down each one of those things, and in each one of those things you can see the number of milliseconds, over three different tries, and where it went.

I live in Salt Lake City, so the first things you see up there are Salt Lake City. If you know your American geography, it then sent it east to Denver, Denver sent it east to Chicago, Chicago sent it east to Washington DC. It rattled around Washington DC for a while, was handed off to Yahoo. The website I was going after was finance.yahoo.com. Goes off to Yahoo. Yahoo went and took it through a series of cryptic things, but I did do some work to find out where these were located, and it sent it all the way back across the United States to right here, Sunnyvale, and that's where the server was that provided the service.

So my traffic went to Washington DC, rattled around for a while, came back to Sunnyvale, and you can see that it took about 65 to 80 milliseconds. The amount of time that people can distinguish that something takes is about 50 milliseconds, so this is a noticeable amount of time just to go and get my request to Yahoo and get it back. That's not providing for any time for the server to do its work of actually fetching the information I'm interested in.

So that's why this thing does not change, even though my end system was much more — I changed the speed there. So going to a gigabit on that last link is probably not going to make that much difference. That gives another experiment. This was done by Microsoft Research. This shows you how long it takes with a bandwidth of 1 megabit. You can see all these are all about the same. When they did it again at 10 megabits, you can see that when they invoked something called SPDY, speedy, which goes back and forth fewer times, that the bottom two lines there are lower and it takes less page load time.

So we think that just from these two graphs the point to take is that as the speed goes up, the importance of the length of that chain goes up. As the speed goes up, the fact that there are more links in the chain makes more difference. So as the internet grows and as it gets faster, the length of that chain makes more difference.

So what's the obvious thing that's going to happen in the future? Make this chain shorter. That's about the only thing we can do. We have speed of light considerations of getting things out there. The number of times it gets handled makes a difference. So we want to go to a local cloud or locavore computing.

How soon will this happen? Well, interestingly enough, it's already happened, starting to happen. AOL has begun putting in what they call micro data centres – this is a micro data centre they put in in Dulles, Virginia – to be closer to the end-user so they can be more responsive. And studies have shown that people buy more and that people will be able to stay with your service longer if you're more responsive. So one of the things that Mike Manos, the Chief Technology Officer of AOL Services, has said is that we expect that micro data centres will enable us to roll out five times the amount of total compute capacity in less than 10% of the cost and the physical footprint of a traditional data centre based deployment. And if you want to go look this up on the web, just Google or Bing, your choice, whatever service you want, micro data centre and AOL, and you'll find this quote and the other pictures.

So this is what's going to happen. It's also happening in research and education in the United States. There's a research project by the name of Genie. And the Genie Project not only facilitates this local cloud computing, but it also facilitates federating these local clouds, bringing the local clouds together so that when you need more compute than one of these local data centres can provide, you can do that by aggregating them together. So you can have the best of both worlds. For simple things, very responsive; it's close to you. For more complex things, you can federate services from multiple local clouds. It will take a bit longer, but still you'll be able to have that ability to aggregate lots of compute power on what you need. And that's the logo for Genie.

Okay. So what I've been talking about is, instead of more bandwidth is better, more responsiveness is better. And to get to more responsiveness, we need to decrease latency, decrease jitter, and a major step to be able to do those two things is to be responsive by having local facilities. There are other ways of reducing and jitter and those are important also, but I don't have time to talk about all of those this morning.

Let me move on to the second big point, is that today we tend to think of, with a big enough pipe, it doesn't matter, I can just toss it all together, and we're going to, I think, end up changing that paradigm. So, first of all, let's go take a look at what I have at my place.

I don't know if you have all of these things at your place, but that laptop right there is the one in the top right. I brought it with me today, but it lives at home. And you can see that I have a couple of desktops, one of which I bought for $200 but still does a lot of interesting things like record my television. I have my own home-built television recorder. Printer. The phone is this phone right here. It's in my pocket. So, all of these things are contending for the bandwidth on my internet connection.
And what happens is the tragedy of the commons, the traditional, typical tragedy of the commons. Today's internet only works well because the utilisation is so low. If we had high utilisation, it wouldn't work well at all. In fact, I have Comcast, as you know, and Comcast says that if you use more than 4% of your capacity on a long-term basis, your usage is excessive and they have the right to go and charge additional money because I'm using excessive bandwidth.

On my phone, which happens to be a Verizon Wireless phone, they begin slowing down my data if I use more than 0.03% of the bandwidth available to my phone. Obviously it costs them more on that basis. Now, that is on a 3G data plan, if you have unlimited data usage, and it is a long-term average. There's a little star footnote there. But the utilisation being low is very important.

Applications cheat. Things like Google Search will go and violate the rules of sharing of the road. It will go and grab more than the bandwidth that they're entitled to. Voice over IP services will try to grab more than they're entitled to. Things that think they are better and want to be more responsive than other applications will grab more than they're entitled to. And there are entire companies, like Bideo and OpenClove, who are based on the proposition of making video work well even when they're having to fight and claw for the bandwidth they need.

So this will eventually come to a head. And the question I would ask is what would data centres do? What would data centres do in this case? And the answer is, of course, what the data centres would do is they would go virtual. They would go and have a virtual server per application. Instead of sharing this and putting them all on one server – we used to call that time sharing – they now go and they have a server per application, a virtual server per application. The server configuration is matched to the application it's running. They allocate these virtual servers as needed, dynamically, based on the load that's being presented. It provides for more fault resilience.

And remember I said, the business model, it's easy to bill the server to the application. Right? You're dedicating a virtual server to an application, so you say hmm, we'regoing to bill this server to the application. And the incremental cost model means that there's more total dollars, because you get to charge every time more work gets done.

So, remember, what would data centres do in this situation we're about to face on the network? They would go virtual. What will happen on the network? I think it will also go virtual. I'm going to change the words here. Did you catch that really quick? I'll go back. Here's the data centre. I'm going to change the word server to network. Instead of a virtual server per application, you have a virtual network per application. You create a virtual network for the application. The network configuration is matched to your application. You allocate as many of these as you need to run your application dynamically, based on the load. It is more fault resilient. It's easy to bill the network to the application and there's incremental cost for service providers. And in fact it might not be the end customer that pays. It might be somebody who is bundling the service.

So I think this is what's going to happen in making that happen. Here's an example of a typical home. Ordinary internet comes in. What else might you want to do? You might want to add another virtual network, maybe some telehealth. Somebody is elderly. They may have a diabetic ulcer which needs to be checked up on, and you'd like to have privacy and you'd like to provision a high-bandwidth, high-quality video link from the home uplink on demand. Maybe you would like have a public safety application, where every home has WiFi that can be accessed by public responders or by citizens who are reporting an emergency and want to beam, for example, video of that emergency.

How about interactive online education? The MOOCs that were talked about yesterday. How can we get high quality, low cost, specifically for that piece of education? Or what if there is something that is very life dependent? What if there is dialysis control, that you need a high-reliability connection? You need to make sure this connection always works, because if not you might be miscontrolling that dialysis machine at home. So these could all be virtual networks using network function virtualisation – I'm sure you will hear that term several times today – to be able to provide multiple services into the same home over the same pipe.

So, again, here's my summary. I think that tomorrow responsiveness, latency is going to be the order of the day. It's going to be provided in part by using lower latency equipment and products. They exist already for the finance industry. I think they're going to come into more widespread use. I think that keeping things local and being a locavore about your computing is going to be important. The cloud comes to you to be more responsive. And we'll end up with multiple, dynamic, virtual pipes which are purpose built and designed to run with the properties of each application.

So, if you would like to hear any more about this next generation infrastructure, I invite you to come join us at an application summit, June 24th to 26th in Chicago. We're going to give demonstrations of applications in education, healthcare, public safety, based on the new technologies I've just talked about, and applications which, as I said earlier, are only possible because of these new technologies being used.

BY Glenn Ricart Ph.D, Founder & CTO, US Ignite