Denis Vilfort: Hi, I'm Denis Vilfort from Hewlett Packard Enterprise. And today, we're going to talk about a very exciting topic, edge computing. You may have heard of other paradigms in the past, such as data center computing, cloud computing. And now, the new paradigm edge computing, where we're actually moving computers to where the data is. A lot of people are confused by that. They say, "Well, wait a minute, where is the edge? I'm familiar with the data center. I'm familiar with cloud computing, but where's the edge?" And the answer is actually everywhere else. Edge is everywhere human beings conduct any form of business. And actually when they do, they create data.

Now, that's an important point there because if the data is generated by human beings creating all this activity, how can we make sense of the data if we have to move it to the data center, or to the cloud? That's actually the biggest problem, that takes time. Moving the data takes time. So, for example, if I'm doing things like x-ray analysis of a patient's x-ray, I can't sit and wait 20 seconds to get the answer back. I have to get the answer right there and then. What we call real time computing. And that's the problem with data center computing, and cloud computing is they're both centralized. Centralized is the enemy of real time computing at the edge.

So now, we have a problem. We have data at the edge, what can we do about it? Well, we have to obviously move the compute to the edge, so we can make sense of the data streams that are coming in, whether it's catching shoplifters, doing quality assurance, or helping doctors being super doctors in being able to diagnose things better. The compute has to be where the data is originating, and where the data needs to be analyzed in real time.

Okay so, now we've got that squared away. What's happening with that data? Well, it's growing by leaps and bounds. Why? Well, because of cameras, sensors, all that stuff. It turns out that roughly 55 billion devices will be deployed at the edge. And not only that, turns out that more than half of the data is actually already at the edge growing quickly to 70% of all data, and all apps, and it doesn't stop there. So, the edge is really where the action is. A very important point.

That means we're beginning to see an evolution away from data center centric computing, and even cloud centric computing. Where now we're beginning to see a tipping point where half of the data, as we've just discussed, is going to be out at the edge in very short order. Now, when that happens, we're beginning to see more, and more data grow at the edge. And the data that is left in the data center and the cloud actually dissipate. If you look at this graph here, you'll notice that, in a few years, we'll actually have reached peak cloud, where a cloud begins to actually decline in the amount of data that it processes. While the edge continues to grow.

So this is an exciting time, so you say, "Well, can you give an example of this?" Yes I can. And it turns out that it is an example where we started using cameras. As you can see in this picture, this is about quality assurance. It's actually a picture from our own manufacturing line, where we're actually manufacturing servers. So one server after another is coming off the assembly line. But here's the challenge, using five cameras and live camera feeds, we only have around one second to determine whether this specific server has been configured correctly according to the customer's order. Because if we don't get it done within one second, it'll actually only be

another second before the next server's coming in. So we need to make the decision right there and then, is this pass or fail?

Now, how's that done? Well, it's done by actually taking around 75 megabytes worth of data that are now being real time analyzed as the data is being seen by the cameras. Like I said, we're seeing it, not just in two dimensions, but with five cameras, 75 megabytes of data needs to be processed in less than a second. If we can do that, then we can cut down incorrectly configured servers that are being shipped, and customers such as yourself will be a lot happier getting exactly what you ordered, but what's not so obvious is that when you do this, you are actually cutting 96 seconds out of every quality assurance step. And when you are building essentially thousands of computers, we're doing in this particular plant 45,000 servers a month, that adds up.

So doing a little math, which I'm not going to bore you with here, actually leads us to a total of 25% additional manufacturing capacity just by shaving a few seconds off the quality assurance step. So, obviously, a lot of people want to do that. And the human eye is our guidepost. A human eye creates a ton of data, and it turns out that we actually, and Samsung last year came out with the 200 megapixel Silicon. And we're quickly approaching the ability for a computer to have better vision than a human eye. With that, we can do things that we couldn't do before. But here's the challenge, guys, those cameras as they get more and more higher resolution, they can see things we couldn't see. We can have better, better quality assurance. We can have better, better healthcare. We can actually protect us against thieves and other things. We can actually have a better quality of life if we can get the data processed.

And as the data that's being generated by these cameras go up exponentially so does the time it takes to move across a network connection. You may have heard that people say we don't have enough latency, or we have too much latency. It's not a latency problem. It's a bandwidth problem. What I can't afford is moving all this data to the cloud, or to the data center for processing. The round trip takes too long because I have not enough bandwidth. So what I do instead, like we just discussed, I'm putting the compute next to where the data stream is. And that's a trend that is growing exponentially.

Now, here's the kicker, it's because I have a real time window that sometimes can be milliseconds, if not microseconds, by which I have to make the analysis, I have to make the decision. If I push my transport time and the data outside that window, I can't get the job done. I'm making compromises. They're basically workloads that are no good, if I have to encounter a data transfer delay. We have to get rid of that delay by putting a computer next to where the data is being generated.

Okay so, you get that. So you say, "Well, does that mean I have to put a whole data center at the edge?" Not at all, because one of the things we have noticed is we are cramming more, and more, and more cores of CPU into a smaller and smaller space. We are actually just a few years away from getting 1292 cores of CPU in just three rack view. You're familiar with the rack view. Just three of those service and a little cluster can do as much as a whole data center used to do just a few years ago. We can replace what is essentially four to three racks with just three little

servers, and put that out at the edge. So, thanks to the miniaturization of the data center, we are now in a situation where we can put the compute out next to the data where it does the most good. Pretty cool.

So, with that, what does that mean for the whole IT planning? Well, turning out that the old data center is actually becoming splintered in a million pieces and being distributed across our distributed enterprise. I was talking to a customer recently that says we're operating 360 manufacturing plants worldwide. And the data needed to do what that plant needs to do is at that location, and it cannot move. Well, if that's the case, the cloud strategy does me no good. A centralized data center does me no good. I need to put a mini data center out in each of those locations. And it turns out that this is precisely what's happening right now. And that is edge computing, where the distributed nest of the compute is taken into our account for the next level of operational excellence. So that's the model we're seeing.

Now, how does that fit in with a cloud first strategy? Well, a lot of cloud vendors started building the cloud and then, subsequently, decided to move to the edge. What we are saying is that's upside down. What you got to do is start with the edge. You start with the edge, and then you say, "Hey, if we can leave it all at the edge, for example, one edge location can copy data to another edge location, and protect our data that way we may not need the cloud." But because many people have a hybrid cloud strategy, the cloud is as secondary concern. We use it for centralized control, centralized backup, and the like. And that is an edge in the strategy. The cloud is secondary to understanding what's happening with the data at the edge.

Okay, you're beginning to see these are some very sexy ideas. How do we get after it? Well, it turns out we at HPE have the platform that does this. It's a platform for your distributed enterprise. It encompasses edges. It encompasses clouds, public clouds and others. It's an edge to cloud platform. So when you look at our platform, what is it made up of, you might ask. Well, I'm not going to get into a detailed technical diagram, but I'm giving you a flavor for it. When we look at this, we have to have a network. It's no good to have a distributed enterprise if there's not a network around. Otherwise, how can you manage your whole estate, if you can't touch it from afar. And certain data, as we just discussed, need to be copied out of the edge location to somewhere else for data protection, and network is great at transporting our data.

Do we stop there? No. We need to have these servers that are doing actual work for us with storage inside. And that's where we store, and compute upon the data. Are we done then? No, we're not. We need a data management layer on top of the hardware, so that we can actually understand that we are part of a distributed enterprise. Things like global name space in the data management layer is a useful thing to have. Are we done then? No, because this is where our applications are running on top of the substrates, both software and hardware. So, now you got that. You go, "Hmm, do I have to manage it all myself?" No, you can actually put machine analytics and AI to work to manage the whole estate, and automate the management of the estate to the tune of doing what you need to get done. So now, this thing can be self-managing, or we can help you manage it. But we will use advanced tools like this as well. And we will manage it from a central cloud location.

Okay, great. Are we done now? No, because as we're expanding the processing through our distributed enterprise, what we now need to do is making sure that security is built in from day one. It cannot be an afterthought. It has to be built in at all layers. And, finally, you don't want to pay for all this. You want to just pay for the outcome. And we need to be able to deliver this as a service. That's precisely what we do at HPE. We actually have Aruba for networking. We certainly have GreenLake to deliver this edge to cloud experience, it's a cloud-like experience that involves all these layers being distributed throughout your entire enterprise. Our service, our data management, our storage, networking the whole bit. And we actually have a security project called Project Aurora that ties the different layers together, and can operate to take corrective action, where it makes the most sense.

For example, a rogue app is misbehaving, well, we can turn off its ethernet port automatically. These are the things that could be done once the whole stack is operated as a platform. And, of course, the ESMO data fabric, that's part of our own technology, that's the data management layer with the global name space. And on top of that, we use the ESMO runtime, and ESMO unified analytics and machine learning operations.

All right so, now you get this. So, how do we take this to market and work with you? Well, depending on whether you are in the financial services arena, or you are in retail, or let's say you are manufacturing, we have vertically aligned solutions where we work with third-party solution completers that run on top of this platform that runs on top of GreenLake. And we take that to you to make sure we can get things done the way you want it done in your way. And then, what's very interesting about this is you say, "My way, what does that mean?" That means you have a choice here of tool chain. If you are a Nutanix shop, for example, we work with Nutanix. If you're a VMware shop, we work with VMware. If you turn out that no, no, we standardize on Red Hat here. No problem, we work with them as well.

So no matter what flavor of ice cream you like in terms of tool chain, we are open. GreenLake is an open thing. And you know, from what working with HPE for years, we've always been open on our servers. We are therefore, open on GreenLake. So, this is where you can have it your way, get a cloud like experience across your distributed enterprise in the flavor of tool chain, or software platform that you enjoy.

So I'll give you a quick little example. Here's a big healthcare provider, operates across 140 locations. How did they do this? Well, there's actually a couple of layer cakes here. First, they provide a decision support, AI driven decision support for doctors and nurses for the purpose of doing diagnosis. So, they have a compute edge that is actually part of a laptop with some GPUs, and they're hooked up to what? To an x-ray camera. So they can look at a patient's live data and discern right there and then with AI filters, if you have a cancer, or any other issues they need to drill into. And, of course, you have multiple of these doctors at a given location. But because you operate worldwide with a GreenLake layer underneath, this particular vendor actually does it in not just one hospital, but many hospitals. So, in each hospital location, there's a one to many relationship with compute running, controlling that hospital's autonomy.

Now, what does that mean? Is there only one hospital in a hospital chain? Of course not. There are many hospitals in a hospital chain. And there is a centralized data center, or a cloud instance that actually fans out across all those in terms of control. Again, this is where networking is so important. I can manage my whole estate as one large fleet of locations, but they are like a span of horses, they're all pulling me in the right direction. And I have control of all of it through the reins. That's what networking gives us.

Does that mean I can't use public cloud as a backstop? Of course not. It might be a good place to put patient records for purpose of backups, and so on. So here we have it, edge to core to cloud. And the stack that's unifying all this is the ESMO data fabric. It runs on the cloud. It runs in the data center, and it runs on all the edge locations. And it interconnects with all the other building blocks that you choose to run there, whether it's containers, virtual machines, or bare metal, doesn't matter. You have it your way, but the data fabric runs seamlessly across all the locations.

Okay so, that's pretty much it. How do you get started, you might ask. This is a lot, I know that I speak fast. We went through it quickly. But here's the idea, we have professional services with already established solutions and blueprints that'll get you going, whether you want to start, and build a smart hospital, smart manufacturing, whatever, our professional services people are here to do a workshop with you, and get you started. So, all we're asking for is consider HPE as your partner when it comes to the new distributed enterprise with GreenLake.