



The Deloitte On Cloud Podcast

David Linthicum, Managing Director, Chief Cloud Strategy Officer, Deloitte Consulting LLP

Title: What's going on at the edge with Carnegie Mellon's Satya and Deloitte's Rahul Bajpai

Description: In this episode, David Linthicum talks with Carnegie Mellon's Dr. Mahadev Satyanarayanan and Deloitte's connected edge leader, Rahul Bajpai, about edge computing. The group discusses use cases for edge, connectivity, and edge clouds—or cloudlets. They also talk about which industries are leveraging edge computing successfully, and how these companies are returning value to the business with their edge deployments. Finally, Satya makes predictions about the future out on the edge.

Duration: 00:27:33

David Linthicum:

Hey, welcome back with this On Cloud podcast. Today on the show I'm joined by Satya, a professor at Carnegie Mellon, and he's been on the show many times but I'm going to have him reintroduce himself, and Rahul Bajpai, a principal at Deloitte. How are you guys doing?

Satya:

Doing very well. Thank you very much for inviting us.

David Linthicum:

Great to be here and I'm looking forward to this topic because I always love topics where I get to learn as much as participate in the conversation. So, before we get going, Satya, could you introduce yourself even though you've been on the podcast a few times? We have a lot of new listeners out there, love to get to re-know you.

Satya:

Certainly. My name is Satya. That's actually short for Mahadev Satyanarayanan. I'm the Jaime Carbonell University Professor of Computer Science at Carnegie Mellon University, and I'm also the very first Cloud Institute fellow at Deloitte, and I'm delighted to be on the show. And my research focus has been on edge computing. I'm viewed as the father of edge computing.

David Linthicum:

Wow, that's awesome. And we're going to hear a lot about edge computing during this talk. Rahul, introduce yourself.

Rahul Bajpai:

Thank you, David. So, Rahul Bajpai. Again, thanks for having me on the podcast. I'm a principal at Deloitte, and I lead our Connected-Edge growth offering within our cloud, sort of as you, our cloud business focusing primarily on edge computing, advanced connectivity, as it pertains to our large enterprises as well as our technology and telecommunication service providers.

David Linthicum:

So, let's just kind of start the show by getting into the basics of edge computing. I know we talked about a lot and did a Knowledge Short on edge computing, but introduce the audience to the concept as well as how it relates to cloud. Satya, I'm going to go to you first.

Satya:

Okay. It's perhaps most useful to ask what is it that edge computing can actually do for you. Why do we need it? And the answer is, while cloud computing is very powerful and obviously is here because of its many strengths, it has certain areas in which it is challenged. So, one is bandwidth scalability. If I have a large number of video cameras transmitting data at high data rates, the collective ingress into the cloud is huge. It's much easier to do the processing close to the point of capture. Second, if after that processing you need to act on the output, the end-to-end latency matters. So, the ability to have a few milliseconds of latency to edge computing rather than 100 milliseconds to cloud makes a difference. There are also data privacy issues where the ability to process the data close to where it's captured, strip it of privacy sensitive information before passing it upstream is valuable.

And there are also data compliance reasons. For example, Europe's GDPR restricts what kind of data can be exported, so the ability to process the data where it's captured can be valuable. And there are occasions and situations in which the cloud is simply inaccessible because of disruption to the connection. And so, again, the computing at the edge can be valuable.

David Linthicum:

That's the most comprehensive explanation I've heard in a long time. Rahul, you have anything to add to that?

Rahul Bajpai:

Yeah, I couldn't agree more with the way Satya characterized it. From our enterprises lens, it's pretty much the same set of challenges, albeit different flavors as you traverse across industries. As an example, some of our large retail enterprises are really looking to solve some of their OT objectives to meet their OKRs on-prem closest to their facilities or their stores, and they have thousands of them across the country. But also looking to retire some of the technology debt they've been carrying for a long time. So, the ability to bring a cloud native architecture on-prem that can still do the modern application processing enabled by computer vision, enabled by these cutting-edge IOT sensors and devices to drive real-time actions with very low latency, high speed, and sometimes processing high amounts of data to derive that end intelligence.

That seems to be front and center. A similar use case would be in remote locations where energy, oil, and gas enterprises are looking to drive better actions in their mining and drilling processes and how they can benefit from harnessing some of that intelligence and some of the applications that can reside on the edge and drive those real-time outcomes for them.

David Linthicum:

Yeah, and one of the confusions that's out there, I think the press kind of facilitated this five, six years ago when kind of edge came on the scene. They go, "Well, it's going to replace cloud computing and they're mutually exclusive," and when the reporters used to talk to me, I said, "No, that can't be the case because it's got to be—number one, it has to be on the edge of something." This is just really a sound architectural option so we're able to leverage technology in a better, more optimized way than forcing everything in a centralized cloud model. So, Satya, what's your take on that and what would your advice be to people who are still confused about that?

Satya:

You are characterizing it very well. The cloud has strengths. There are certain strengths it's very important not to lose sight of. First of all, data centers are like fortresses. It is the most secure place to put your data to perform a computation, and the cost per unit of compute is the lowest that it will be anywhere because the economies of scale drive down the per-unit cost. So, in terms of longevity of data, if you need to put data in a place that you are sure will be there five years from now, ten years from now, will be backed up, will be secure and so on, where you can execute applications which do not have any of the constraints that we just discussed, that both Rahul and I talked about a moment ago, the cloud is the right place to do it.

So, the goal is not to replace cloud computing by edge computing, but it is to recognize that they both have extremely important roles to play and to leverage their strengths to create end-to-end architectures that seamlessly combine their strengths from the viewpoint of the user at the edge or the cyber or physical system at the edge.

David Linthicum:

Yeah, and I think this is people getting more pragmatic about the use of technology and also the cost of some of the on-premise technology in the last ten years has dropped significantly, and your ability to leverage architectural options to get to return more value back to the business is going to be very important. Let's talk about edge computing in terms of connectivity. So, obviously, connectivity is everything in terms of edge computing because we're in instance leveraging edge computing because of the connectivity advantages. So, Satya, I'm going to go back to you for this and, Rahul, then I'll go to you. What are some of the enabling technologies that enterprises are leveraging to do advanced connectivity and edge computing these days?

Satya:

The emergency of 5G is very well positioned to benefit and leverage edge computing. 4G LTE is also good, but the latencies and the bandwidths of 5G are significantly better. If you flip it around, 5G without edge computing is actually not that useful. What are you going to be able to do with it that you couldn't otherwise with 4G? Perhaps deliver more video, perhaps deliver it a bit faster, but that's it. But if you combine 5G and edge computing, you can create cyber-human and cyber-physical systems that you could not create in any other way, and those will have transformative effects for industry and enterprises.

David Linthicum:

So, Rahul, do you see 5G as really kind of a game changer in terms of our ability to use IOT's evolutionary technology or evolutionary technology? And I'd love to hear how enterprises are leveraging it since you're close to the customers.

Rahul Bajpai:

Yes, absolutely, David. And in fact, the way we see 5G as an enabler to a lot of the connected edge ecosystems, there are enterprises that are trying to create for their benefit. But 5G comes in different forms depending on the use case or the archetype we're talking about. In a true mobility environment

where you have autonomous driving or like an ambulance on the go that has a better connectivity need into the healthcare system, absolutely commercial 5G has a big role to play.

As you come in closer to the enterprise and inside the enterprise, there are use cases that can be fueled by private 5G, leveraging our CVRS spectrum, or in some cases, an advanced connectivity mechanism like Wi-Fi 6 can also do the trick. Of course, the key thing to note here, as Satya said, is that advanced connectivity, very high throughput, very low latency, and the ability to connect several hundreds and thousands of sensors, devices, and endpoints which are graphic user—like computer vision type interfaces or even regular sensors, that does require a very highly advanced connected network. 5G absolutely is here to stay, but alternatives like Wi-Fi 6 and private 5G will fuel a lot of our enterprises as well.

David Linthicum:

Yeah, we really are moving to ubiquitous computing ultimately where we're going to leverage the data and the processing where it exists because connectivity is becoming much, much better than it was just a few years ago. So, Satya, I'm going to go to you. Let's talk about the trends including the use of edge clouds and basically that growth of the area. When I talk about cloud computing, I can't leave out discussion in terms of the architectural option and people using edge clouds, which are the ability to leverage clouds in different ways, but you explain it to us. Why would I want to leverage an edge cloud, what is it, and how are enterprises looking to leverage it today?

Satya:

Sure. I have used the word cloudlet to mean edge cloud, so I'll continue to use it. They mean the same thing, but the idea is concentrations of computing close to the edge using technologies that were originally created and conceived in the cloud. So, in the course of creating and developing cloud technologies over a decade from about 2006-07 to about 2015, even through today, there have been a whole host of innovations, including for example, hardware innovations for use inside the data center, software innovations and software structuring innovations such as the concept of cloud native structure, the idea that you can do scale out, the ability to cleanly separate state, stateful, and stateless components. All of these things are technologies that have come out of cloud computing. Every single one of them is valuable at the edge.

The only difference is, instead of applying them in a giant exascale data center which spans a football field or larger, it is being used in a mini data center, much smaller, perhaps it fits in a small room, whatever, it's on-prem perhaps, or it could be onboard an aircraft or a ship. So, the vision that edge computing is not about replacing the edge, but it's about replacing the cloud but it's about bringing the cloud closer. That's a metaphor that I find very helpful in explaining edge computing and its relationship to cloud computing.

David Linthicum:

Yeah, it's also valuable options in terms of the ability to leverage the same skill sets and leverage the same software systems, cloud services in kind of a private cloud deployment and edge based deployment, which provides you with flexibility, so therefore you can keep the same skill sets, keep the same architects around, and examples of this would be that are really kind of building things that exist not only in the public cloud but also in the private cloud instance which allows enterprises to take advantage of both architectural advantages of doing both, and there's tradeoffs in leveraging things that are in a public cloud or within your data center. So, Rahul, what are some of the use cases here? Why would I want to use a cloudlet and why would this be some kind of a good option for enterprises out there that are trying to make their applications scale and function better?

Rahul Bajpai:

And like Satya said, David, cloudlet or edge compute architectures are helpful for a couple of key reasons. So, one is if you start looking at some of the net new use cases, think of universities trying to achieve better security on their campuses and they want to be able to track panoramas with 4K cameras. Or think of a high-end car assembly manufacturing line that really wants to apply 4K camera computer vision again to assess the thickness of a paint coat on the body of the car and then make corrections in near real time and create a feedback loop that basically allows them to reduce the human intervention or labor involved. If you look at some of those use cases and then couple them with some of the upcoming sort of like focus on predictive analytics using AI/ML or GenAI to improve productivity and drive automation and outcomes.

All of this requires richness of data, which in turn requires the connected applications on a common infrastructure, common architecture, common management tools and systems so that you don't end up with a data swamp. You actually have a data lake to produce those rich insights from, and that's why those cloudlets, or micro DC, or small cloud sort of architectures on the premise would really help our enterprises harness sort of that intelligence that they need to drive outcomes. So, for us, it's a technology that enables larger outcomes of the stack, or of the value chain.

David Linthicum:

So, Satya, we're talking about this as very homogenous kind of relationships or more tightly coupled cloud with edge cloud, the ability to leverage the same brand cloud, for instance, either both on the edge and in the public cloud space. What about heterogeneous edge cloud? What are some of the options out there that people are leveraging? What are some of the architectural challenges and also opportunities to make something that's going to be more cost effective for the business by leveraging cloudlets in conjunction with a public cloud provider?

Satya:

Absolutely. And in fact, if you look at my recent publications, you will see a paper there that just appeared a few months ago called "Sinfonia: Cross Tier Orchestration." And this is precisely the idea. The idea of cloudlets or edge clouds, whatever word you want to use for them—I'll continue to use cloudlet since that's what I'm familiar with—the idea of cloudlets being brought up on demand just in time is a very powerful concept. Imagine a disaster area or imagine a construction site or imagine a military operation. There are many situations in which you wish to create in a very short amount of time, a few hours, a mini data center with 5G coverage of a remote area. It might even be an urban area except the infrastructure is damaged, or it might be a manufacturing plant for some reason which is suffering from outages.

You can think of a long list of use cases, but the point here is that there is no need for the entire chain from edge to cloud to be homogenous. If you can bring into existence on demand edge components and make them work seamlessly with the cloud, obviously with the right level of authentication and

security achieved through the right mechanisms, you have now transformed the kind of architecture that is available for use, and I think that is a very powerful vision. So, the concept of heterogeneous systems which are able to combine devices at the edge, a diversity of cloudlets in the middle, some possibly just in time, and the large exascale data centers in the cloud, that vision is a very, very powerful vision.

David Linthicum:

Absolutely. So, Rahul, go to you. What are some of the industries that are leveraging this technology and how are they leveraging it and how are they finding value with leveraging this technology?

Rahul Bajpai:

I think a couple of examples that I gave across retail, automotive, healthcare, and then manufacturing at large, which obviously cuts across multiple industries, have begun to leverage this technology. But to Satya's point, there is a myriad of options out there that basically allows our enterprises to leverage the best tech that they can for the lowest price point that they are able to. For that, it's a great option to leverage some of the hyperscalers.

Each of them has their own distributed cloud edge offerings. But you can also look at alternative architectures where bare metal with a virtualization stack on top with an orchestration stack on top can still bring that micro data center/cloudlet-like technology and architecture to bear, which some of our enterprises are using today to minimize they want to be able to leverage the hyperscaler technology for the large cloud instances that they need, but to minimize some of the—I would say getting rid of the legacy systems immediately and increase that tech debt or that burden. They would want to wait until they have depreciated assets over time, so they are trying to leverage existing assets, which is making them look at alternate architectures with the open-source technology.

David Linthicum:

So, what are some of the value points that businesses should consider? What's the elevator speech in terms of this is how you're going to make money by using this technology and this is an instance or an example of where it's made money?

Rahul Bajpai:

Absolutely, David. So, I think the way we have seen our clients embark on this journey and being with them over the last 24 months or so, we have obviously seen some hits and misses, but the things that have worked are the following. So, number one, have a clear sort of business case and outcome-based approach of what you're trying to increase in terms of revenue and what you're trying to save in terms of operational costs. I think those two factors play a key role in how you design the architecture and what set of applications you invoke first or leverage first to achieve those outcomes. So, that's number one.

Number two, a very quick but I would say business-case led-assessment of what the technology stack is today, what's your plan to retire, and how you're bringing on these new architectures to drive those outcomes, either to increase revenue or reduce costs. I think that should have a very tight linkage.

Third, which we have seen work the most, is really trying not to boil the ocean but taking a very small area of the business, prototyping that, doing a pilot, doing it for about a three-month duration, looking at what the outcomes are relative to the costs and the effort required, and fine tuning both the technology stack and the operational processes that go around them to ensure that you are getting the biggest bang for your buck and not having to reinvent the wheel for every other use case that you have to launch or every other business area that you have to cover with this technology enablement.

David Linthicum:

Those are great data points. So, Satya, I'm going to pick on you a bit because you did this last time, you did a great job. A lot of the predictions you made came true. Say we're doing this podcast in five years. What do you think we're going to be talking about in terms of edge computing and its use within the cloud computing marketplace and what technologies do you envision evolving to be very substantial within a five-year period of time?

Satya:

So, the way I would like to characterize that is by characterizing a class of new applications which are just starting to emerge. I call them edge native applications. Everyone is familiar with cloud native, but edge native means something different. Edge native applications are applications that are critically dependent on edge computing, on cloudlets. They cannot run by themselves on mobile devices or wearable devices. They cannot run purely in the cloud, and the reasons might be one of the reasons we gave just a few minutes ago. Maybe the bandwidth needed is too high or the latency is too much or so on. And yet the combination is value to the end user that is simply not achievable today without edge computing.

Five years from now, I would love on a show like this to be highlighting the top five or top ten of those edge native applications and saying, you know, this is what edge computing can do for you. It frees you from the constraints of having to go to the cloud. At the same time, it gives you cloudlike capabilities at much lower latency and higher bandwidth to access, and as a result, you're able to do the following things that you could not do otherwise. Some of these are already starting to emerge. Augmented reality everyone has heard of. AI and ChatGPT is something else you have heard of. Imagine the power of the two being combined.

How do you place something like the power of ChatGPT into the inner loop of an augmented reality system? Could I create a pair of wearable glasses that I look around, I see you, I've forgotten your name, but it whispers in my ear who you are and where I last met you and what your wife's name is and so on. The notion that we have applications of this kind which could be usable, and this is just one example, it's kind of science fiction-ish, but every time I've described it, it becomes less science fiction as we're getting closer to achievability. So, that's just one example. In drone control, today it is the case that drones are in the media, they're a lot of hype, but there are important use cases that nobody is talking about which are actually going to be where the monetization comes from.

If you think about the construction industry, you think of New York or Chicago and inspecting the outside of a building, putting up scaffolding, those are all very expensive things. Those are things that small drones, lightweight drones which are safe to fly within city limits in dense urban areas could do the

inspection very carefully and very well, combining autonomous drone flight with computer vision and machine learning to do the actual analysis of the examination—roofs, the skins of the buildings, and so on.

The ability to use 5G and ground-based cloudlets effectively allows you to have brilliant drones that are small and lightweight, and that will be another game changer in many areas.

So, I've just given you a few examples, but we need to start moving away from thinking in terms of what can I move from the cloud to the edge and really start thinking about what we can do that we could never do before, and that's where the sweet spots and the huge wins for edge computing are going to come from.

David Linthicum:

I love that answer and I love the predictions you provided. I love edge native. That's probably something that's going to be a buzzword—you heard it here first—we're going to see a lot of and probably when you're writing. So, Rahul, where can we find more about you and your services and your team on the web, and how can the listeners learn more about edge computing in general from Deloitte resources?

Rahul Bajpai:

Yes, David. So, you know, on our website there's obviously a storefront, I would say, of our broader cloud business offering, within cloud what we do within connected edge, advanced connectivity and edge computing focused on 5G with some whitepapers and POVs published around our experiences with different enterprises and how they leverage. So, a lot of information, obviously, is on our website. But our practice starting with me are always available to address any questions from our clients or any interested colleagues from the broader professional network, and happy to kind of tune in and see what we can do and how we can help.

David Linthicum:

Great. Satya, same question. Where can we find more of your research on the web?

Satya:

A full publication list is at my website, www.cs.cmu.edu/~satya. So, you can find me there, and it's got a publication list and everything I've written and pointers, too, including describing edge native applications.

David Linthicum:

That's great. I'm looking forward to keeping up with this and getting you back on the podcast at some point in time, Satya. It's always great, very informative. So, if you enjoyed this podcast, make sure to like us, rate us, and subscribe. You can also check out our past episodes, including those hosted by my good friend, Mike Kavis. Find out more at deloittecloudpodcast.com. If you'd like to contact me directly, you can email me at dlinthicum@deloitte.com. So, until next time, best of luck with your cloud journey. Everybody stay safe. Cheers.

Operator:

This podcast is produced by Deloitte. The views and opinions expressed by podcast speakers and guests are solely their own and do not reflect the opinions of Deloitte. This podcast provides general information only and is not intended to constitute advice or services of any kind. For additional information about Deloitte, go to Deloitte.com/about.

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor.

Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

Visit the On Cloud library
www.deloitte.com/us/cloud-podcast

About Deloitte

As used in this podcast, "Deloitte" means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

Please see www.deloitte.com/about to learn more about our global network of member firms. Copyright © 2024 Deloitte Development LLC. All rights reserved.

