Building Up The ‘Edge’ for A Responsible Society

Nitin Dahad (00:05):

This is the Smarter World Podcast, focusing on the technology and issues behind today's connected world. I'm host, Nitin Dahad, technology journalist and editor at EE Times and embedded.com.

(00:17):

In this episode, we'll discuss the role edge processing can play to help make the world greener, more productive and safer. I'll be talking to Gowri Chindalore who heads strategy of edge processing at NXP and Amanda McGregor who leads Edge processor innovation at NXP.

Gowri Chindalore (00:33):

Hello Nitin.

Amanda McGregor (00:33):

Hi Nitin. Great to be here.

Nitin Dahad (00:35):

I did a recent podcast, I spoke to Ron Martino, who leads the NXP edge processing business, and Ron made it very clear that the benefits of edge processing go beyond convenience. Gowri, in your role as head of edge strategy, how do you see edge processing benefiting society?

Gowri Chindalore (00:52):

About edge processing, Ron is right, it's more than mere convenience, it's more than being able to turn on the lights in a house with a voice command. It can actually help us build a better society, a more socially responsible society. What do I mean by that? It means edge processing can help us reduce energy consumption as a society, reduce the energy consumption, which means a green environment. It can help us be safe and protect personal privacy. And it can help reduce waste by improving productivity.

(01:23):

But we got to be careful with this. None of these are automatic. There are going to be tens of billions of devices around the world that are deployed, if we don't do it right, we could end up having exactly the opposite effect. But if done right, edge processing can help us realize a society that is green, safe and productive.

Nitin Dahad (01:43):

You did focus very well on being green, but in that context of edge processing, what does green actually entail? And maybe give me some examples.

Gowri Chindalore (01:50):

Green in this context is using as little energy as possible to perform a specified task. In other words, do more with less. Let's take the example of a video doorbell that can recognize the members of the family and when they come to the door, it recognizes the face and it unlocks the door. That's totally convenient and really cool.

(02:11):

But there's two ways that this thing can happen behind the scenes. In the first way, the image is taken by the doorbell and that image is transmitted to the cloud or some other entity and it waits for the decision to come back on whether to unlock the door or not. The second way is the door bell takes that image, processes the face locally and then makes a decision whether to unlock the door or not.

(02:36):

The difference between these two is the energy consumed in doing so. Just imagine a five megabyte of a JPG image that's sent to the cloud about 20 times a day when people come in and out and processing that in the cloud and coming back, waiting for the decision to come, that's about 10,000 times or more energy expensive compared to taking that image, processing it locally and making that decision. So, an edge processing that actually has the capability to process and make an intelligent decision is several orders of magnitude green and it saves monumental waste of energy.

Nitin Dahad (03:16):

Yes, and that really sounds like a no-brainer, I guess, you're doing a lot on the device itself and then, you have a lot of intelligence in terms of the embedded software that allow you to do that. Did you want to expand on that or maybe Amanda wants to expand on that a little bit?

Amanda McGregor (03:30):

Yeah, Nitin, I'd love to expand on that. This concept of reducing monumental waste of energy is just so important and it's really the foundation of how edge processing can really give you that ability to go green. But as you mentioned, we really need to take it a step further than just having the ability to give you green. We need to make it greener. There are really clear benefits to reducing the edge to cloud energy waste, Gowri just gave a great example, but we have to think about what we're actually doing with the processing at the edge, that's what really matters here. And I'll give you some examples.

(04:04):

So for example, an edge device, it's always on and running in a high power state. Even an edge device that is in a low power state yet consumes significant vampire power, that's not green or greener either. You really need to have these edge devices take it one step further and really be thoughtful about how they can be truly efficient with resources, how they're really becoming more aware and creating different states of awareness, the amount of on time, the leakage power. And we really see as there are many more government regulations being introduced

such as, one example, the European Energy Commission, really driving more requirements around reducing energy footprint and I think consumers are becoming much more aware of that as well. So you have regulations coming into play that are mandating more tighter restrictions, you have consumers becoming much more aware of their energy footprint and I think that's really driving this technology shift towards enabling a more responsible edge.

Nitin Dahad (05:01):

I was just going to ask you on that specific aspect of reducing energy footprint, what sort of order of magnitude do you think you can achieve in terms of improvements doing things at the edge versus doing cloud? But also, are there any other techniques you can use to reduce that even at the edge?

Gowri Chindalore (05:17):

We recently did a calculation, we wanted to understand how much of energy it takes to do some calculations at the processor itself, at the edge, was sending the data to the cloud. And we did an estimate with what can be done with a hundred picojoules of energy. Our calculation said to transmit bit through a [inaudible 00:05:37] network or a 5G, a hundred picojoules can transmit only 0.0002 bits, versus, an MCU uses the same energy to do more than hundred floating point calculations, 2000th of a bit with hundred picojoules versus doing hundreds of calculations at the CPU. So that gives you a feel for the orders of magnitude difference in the energy required.

Nitin Dahad (06:02):

That's quite a difference. Amanda, how is NXP doing things differently at the edge to improve that energy consumption?

Amanda McGregor (06:08):

I'm going to come at it from more highly capable applications' processor perspective. We know there's a lot more functionality being driven into these edge devices and so, that inherently requires more processing capability, more ability to connect to different sensors and inputs. So the way we think about doing things differently is, clearly, we need to provide that functionality into these devices, but that capability isn't always needed all the time. And so, where we can move into states that are traditionally more of a micro control or lower power state and where there still some level activity, right, you maybe still need to be picking up on some activity through sensors, maybe you are waiting for voice activation, being really intelligent in the way that we manage the resources both on the chip and both externally to really make sure that we are just doing the absolute minimum at that point and really switching off all of the other capabilities on the chip. And so, that becomes really important to best use the resources and provide that scalability from very low power all the way up to more capability when it's needed.

Nitin Dahad (07:13):

Essentially, it'll wake up on triggers and those triggers might be voice activation or it might be something else.

Amanda McGregor (07:18):

It could be movement, it could be listening for other sounds. There are many different ways we can put the product into a state that it can still be on, still be aware, still be listening, but doesn't need to be consuming as much of the resources.

Nitin Dahad (07:31):

Let's move on to something about safety now. Technology obviously unlocks tremendous opportunities to improve safety in our homes and at work. What kind of safety transformations can we expect in our daily lives?

Amanda McGregor (07:42):

Yeah, Nitin, there are so many great examples to share on this topic. I think we could take this entire podcast and just turn it into examples, but there are some obvious examples, but one I really like to talk about, and maybe it doesn't get as much awareness, is how edge processing technology can really improve the safety of workers who might be in more of an exposed environment. Think of a construction worker. One statistic we have from the Bureau of Labor Statistics reports that only 16% of workers who sustain head injuries on the job wore hardhats even though many are required to. So we see innovation from companies who are developing smart hardhat technology, so there could be real-time feedback on who's wearing them and who isn't, so that gives monitoring and tracking capability.

Nitin Dahad (08:29):

And would that be something like a movement sensor or [inaudible 00:08:32] sector or something like that?

Amanda McGregor (08:32):

Yeah, exactly. It could be. And then, what if you could take it one step forward and maybe send warnings and alerts in case there's danger? Maybe these smart hardhats have some kind of vision capability in them and then you start to get both the feedback from the worker wearing them, so that's ensuring that they are being safe on the job, but then adding further to that and providing more environmental feedback to actually warn the person who's doing the work if they're at risk. So I think there's a lot of really good ways of getting feedback that can really provide dramatic improvements in safety.

Nitin Dahad (09:04):

That picture you paint gives me the image of a minor down at a shaft here, and this is very old-fashioned, but having all this hardhat which can have all the environmental sensors, impact

sensors and everything else and you just enable people to understand what kind of danger he's in.

Amanda McGregor (09:18):

And then, you could even extend that to recreational use and maybe you are exploring a cave or doing something that's a little bit more of an extreme sport and being able to use that to your advantage so you can also have that enjoyment but be operating in a safer environment too.

Gowri Chindalore (09:32):

This conversation between you and Amanda reminded me of an interesting statistic that I had recently read involving home security systems. Apparently, majority of homeowners, like 70% or so, don't use the alarm systems because it is really a chore to turn it off and turn it on when they leave the house and then come back and enter the code and disarm it. Yeah, granted, these days there's a better case like with a key fob or using your smartphone to arm and disarm the system, but it's still human intervention.

(10:04):

But imagine an intelligent smart home which can actually detect when the last person has left the house, locks the door automatically and arms the system and detects when the first family member has come back and unlocks the door and disarms the system. It's convenient and it is safer and it's using a facility that we already have.

(10:27):

Of course, what this needs is a smart home system that has intelligent devices that are able to communicate with each other and make meaningful decisions, but it's possible and if we do so it really improves even at home level safety and security.

Nitin Dahad (10:41):

I've seen that image of your CTO demonstrating that at various trade shows-

Gowri Chindalore (10:45):

Yes, that's right.

Nitin Dahad (10:47):

... where you can walk into... Using ultra wide band or various other capabilities to enable secure access. So I think that's actually secure but also seamless, I think that's what his point is and your point is.

Gowri Chindalore (10:58):

Absolutely. But it also needs sensors inside the house to detect that a human has left, and the pets are still inside the house. It needs to be able to delineate between the two things, and it's possible, we have the technology to do that.

Nitin Dahad (11:10):

Let's go into data and privacy. We're starting to see more countries adopt new regulations to protect consumer data, we already have it here in Europe with GDPR. Technology and specifically, connected edge devices generate enormous amounts of personal data, how is the industry preparing for this shift?

Gowri Chindalore (11:26):

There's definitely increasing concern around protection of consumer data and data privacy, and rightfully so. Data privacy is the single biggest concern that the consumers cite as the reason they're worried about the expansion of internet of things, and edge processing is actually the first step in addressing those concerns because it's, by definition, processing where the data is generated rather than sending it to some remote location where it could be vulnerable. But that's not enough. Actually, we need to go one step further.

(11:57):

It's not enough to do dumb processing, processing a video image and then not knowing what to do with it, it's not good enough. So we need to add machine learning and intelligence to the edge so that it not only processes the data but also makes a decision based on the information that it just collected, making sure that all the data and processing and information is stored within that local system.

(12:22):

And that's where a trustable system comes into play. Talking to a device in a natural language, that device being able to recognize your face, opening the door, they're all great, but how do we know that this data is not being transmitted to some bad actor through some backdoor? In fact, a recent survey said 63% of the people find connected devices to be creepy and 75% don't trust the way their data is being shared by those devices. What we need to do is build intelligent edge device systems that are trustable.

Amanda McGregor (12:59):

Yeah, I'd love to add to what you said, just really add to this concept of edge devices needing to be trustable because we know that security is really only as strong as the weakest link. And really, there's a whole development stage of product as well as deployment where security can be either enforced or compromised. Nitin, you mentioned GDPR and we know that there are other regulations coming into play as well, I think there's the California Privacy Regulations Act as well and they're going to start enforcing that from 2023.

Nitin Dahad (13:26):

Is that SB-327?

Amanda McGregor (13:27):

I believe it is. And I think from, really, an OEM standpoint, that's going to really motivate the need to seek out technology that can enable them to build a secured product. Security is not something you can just add on after the fact, it really needs to be designed from the ground up, and the ground up starts with the chip itself and then really extends to the entire system and the ecosystem.

Nitin Dahad (13:49):

This is one of the biggest areas that I've been writing quite a lot about in terms of security and privacy, that is actually both the biggest benefit for processing at the edge and doing a lot of intelligence at the edge. But at the same time, you need to make sure that you can have confidence in that ability to trust the data flows within that. Is that right?

Amanda McGregor (14:09):

Absolutely. And there are certifications that can be attained to assist with building the confidence around the technology itself, CCIP, like FIPS, there's different certifications in different regions. But to be able to meet those certification levels, you really need to have that on-ship capability, secure enclave level of approach on chip, and that really provides that whole security in terms of booting software, I mean, zeroing it as a trusted device, cryptographic acceleration, runtime attestation to ensure that the device is authenticated to run on the network, to run that software, that it's not been tampered with, that it's not been cloned, that there are mechanisms in the hardware itself to really ensure that you can meet these certifications and that you can provide and build a trusted system based on that technology.

Nitin Dahad (14:58):

I think we should also make it clear that this is not a one-time thing because it's about lifecycle management and I think you're among several companies that have launched those lifecycle management solutions over the last year or so.

Amanda McGregor (15:10):

You're correct. You put an IOT device out in the field and some may be in the field for three years, some may be seven, some may be more than 10 years and so, you have to think about that whole security lifecycle. We don't know now what will be the security situation in 10 years from now, what's required, and so, you really need to build that true hardware capability that then have the ability to scale and address and update throughout that lifetime of the product, and that's where software and services and cloud comes into play to really ensure that trustability through its lifetime.

Nitin Dahad (15:42):

Okay. The last part of my opening to this section was how is the industry preparing for this shift in terms of enabling data privacy and security? So we talked about the hows and whys, but how's industry preparing for that? How are you seeing with your customers and some of the things that are happening within the industry itself?

Gowri Chindalore (16:00):

So then, to that, in terms of industry preparation, there's a couple of things. One is like you guys were talking about, device management, having the hardware capability present so that it is somewhat future-proofed so that any kind of a future security vulnerability can be patched through software updates rather than recalling that entire device, so that software updateable, putting the right hooks into the hardware is a key. So we are constantly thinking ahead and saying, okay, what are the different hooks that need to be put in place so that a software could patch any kind of an issue in the future? So that's one shift.

(16:38):

The second shift is really creating device management, trusted ecosystem. We announced a family of products that we are going to be launching with our Microsoft partnership. They're going to bring the decades of experience in device management to the edge and they will be able to send software updates to patch any kind of a vulnerability for the devices in the field. It doesn't have to be that, there could be other companies who will be looking into this kind of over-the-air update capability to manage.

(17:09):

And ultimately, it's about decommissioning too because a device, if it's not being used within a network, it becomes the weakest link like Amanda said. So being able to safely decommission that from a network is as important as safely commissioning it. So the industry is-

Nitin Dahad (17:27):

It's like leaving a landmine.

Gowri Chindalore (17:28):

Exactly. Exactly. Yeah. Undetonated mine is more dangerous than... Yeah, absolutely. It is like that. So we are working with a whole set of ecosystem partners and providers to create a complete cloud to edge security that's sustainable over the lifetime of the device.

Nitin Dahad (17:45):

How can we put machine learning to use in the real world for the edge? Because it sounds like something you do in something very compute intensive, but I guess, you can do that now quite well at the edge. But tell us a little bit about how we can do that.

Amanda McGregor (17:57):

Yeah, Nitin, when I think about one of the, probably most well understood real life applications of machine learning, it's really based around vision and voice. But let's go with vision because I think when you think about the ways it can be applied, there are concrete proof points today and then, where you can take that further is really infinite.

(18:17):

So when I talk about vision-based ML, I'm talking about a device that can essentially see its environment or it could take in an image through some other means, but essentially, it's having a field of view of the environment. And so, if you think about how that can be applied, there's a smart door lock, that's a great example and I think it ties into more safe, secure home or office building. Think about only those who are authorized to enter at specific times, their face can be detected, recognized, if they're approved to enter, then they can enter. So I think that's a really good, safe, secure application of ML.

(18:51):

If you think about the medical industry, it's always innovating, technology is really significant in the ways that they innovate. And the idea of having remote medicine, being able to bring more diagnostics to areas that don't have reliable connectivity or even any connectivity and really being able to provide more capabilities to those areas, I think, is really significant in making a difference to people.

Nitin Dahad (19:15):

Okay. I do get that the vision and being able to do that, there's so much data involved, there's so much bandwidth required if you had to do it elsewhere. So I think that's quite important. The next evolution for edge processing, can you tell us a little bit about that?

Gowri Chindalore (19:25):

Let's think about how far we have come. We had mechanical devices decades ago, we converted them into smart devices, then we connected them together to create the internet of everything, then we made them smarter by adding a lot of processing promise to create the edge. What's happening now are the examples that you and Amanda were talking just now about creating an intelligent edge by adding that machine learning and inference capability inside them. From our perspective, the next evolution of this is taking this individually intelligent devices and connecting them into a network that can collaborate with each other to create a collaborative intelligence.

(20:08):

We got to pause on the thought process a little bit, right? It's collaborative intelligence, by that, what we mean is these intelligent devices are interacting with each other, they know the environment they're in and they know the context of some event happening and then, they make a decision based on that context. We call this “aware edge” as in like, it's aware of what it's supposed to do, where it is and what it is going to do. And it's all going to be decision-making by this aware edge network independent of the humans, but based on some of the guiding

principles that we value as a society around ethics and morality. So the aware edge is what we see is the next evolution of this intelligent edge

Nitin Dahad (20:54):

And how far are we on that path to aware edge?

Amanda McGregor (20:56):

Our customers putting into practice these ideas, in particular, on some of our newer products that are really bringing this capability of building out an aware edge. We've recently announced production of products that have hardware acceleration for machine learning, so that's where you can really take advantage of these products to build out this aware edge. So we do see applications where this type of vision processing, this sensing is all being connected and being used in many different ways, whether it's for managing fleets, and by fleets I mean trucks, vehicles, whether it's managing home security or entry, so there are many, many examples, vision in industrial applications and manufacturing lines. And so, I think we are seeing real world examples of these being put into play now and it's really exciting, and the opportunities are accelerating because we're building the technology that can bring this vision to reality. It comes with challenges though, of course, so-

Nitin Dahad (21:52):

I was going to ask you, what are the challenges for aware edge to take shape?

Amanda McGregor (21:55):

I think you mentioned earlier, Nitin, about the amount of processing that's traditionally been required from a cloud perspective to put ML into practice. In deploying this cloud to edge concept of what we can practically do at the edge with the resources that we have, because we have to be more efficient because a number of devices are much greater, so we have to be more efficient in the way we do things, we're closer to the action of what's actually happening, so that helps as well.

(22:19):

But the challenges are really, how do we deploy efficient ML on these devices? Because many of the models, they still need to be trained in the cloud. So you need to be able to optimize and deploy them to actually run at the edge, so that's a challenge that is technology, it's tools, it's ecosystem, it's really multiple factors at play here and multiple areas that need to collaborate together to make it the reality that we see it can be.

Gowri Chindalore (22:45):

So Nitin, I want to add to that last point that Amanda mentioned, right? Creating a collaborative aware edge, if the devices are coming from the same manufacturer, their software could be written in a way that they interact with each other, there's a unique identifying capability and

they trust each other. But we seldom buy products from the same manufacturer, we buy from several different companies. It's important that the ecosystem and regulations and standardizations be built to make this aware edge a reality where different products from different companies and different players could interact with a certain set of guiding principle. So there is a regulation, standardization and a software challenge that's needed above the level of the hardware and devices to make this aware edge a reality.

Nitin Dahad (23:32):

I think on that note, I'd like to close this podcast. Gowri and Amanda, thank you very much.

Gowri Chindalore (23:37):

Thank you very much, Nitin. It was pleasure talking to you.

Amanda McGregor (23:39):

Thank you, Nitin. It was a pleasure talking to you.

Nitin Dahad (23:42):

This has been the Smarter World Podcast with me, Nitin Dahad. Thanks for listening. See you next time.