Intel

Intelligent Orchestration and Management of 5G Edge Services

CORPORATE PARTICIPANTS

Lilian Veras
Intel

Alla Goldner
Amdocs – Director of Technology, Strategy & Standardization

Francesc Guim
Intel CTO Office NEX – Principal Architect and Edge Chief

PRESENTATION

Lilian Veras

Welcome, everyone, to the Intel Network Builders webinar program. Thank you for taking the time to join us today for our presentation titled “Intelligent Orchestration and Management of 5G Edge Services”.

Before we get started, I want to point out some of the features of the BrightTALK tool that may improve your experience.

There's a Questions tab below your viewer. I encourage our live audience to please ask questions at any time. Our presenters will hold answering them until the end of the presentation.

Below your viewing screen, you will also find an Attachments tab with additional documentation and reference materials, including a number of websites and documents mentioned in this presentation.

Finally, at the end of the presentation, please take the time to provide feedback using the Rating tab. We value your thoughts and we'll use the information to improve our future webinars.

Intel Network Builders webinar series takes place live twice a month, so check the channel to see what's upcoming and access our growing library of recorded content.

In addition to the resources you see here from our partners, we also offer a comprehensive NFV and SDN training program through Intel Network Builders University. You can find a link to this program in the Attachments tab, as well as a link to the Intel Network Builders Newsletter.

Intel Network Builders’ partners have been working to accelerate network innovation by optimizing their solutions on Intel technologies. These industry leaders are recognized in our Winners’ Circle program, and Amdocs is a Titanium Partner. Learn more about our INB Winners’ Circle program by clicking on the link in the Attachments tab.

Today we're pleased to welcome Alla Goldner from Amdocs and Cesc Guim from Intel.

Alla Goldner leads 5G and ONAP activities and internal decision-making at Amdocs where her responsibilities include defining and implementing Amdocs’ ONAP strategy, working directly with ONAP members from service providers, vendors and partners, and leading Amdocs’ internal Technical Steering Committee. Her current focus is 5G, Zero-Touch operations, NFV and SDN network evolution, and in particular, the technology, strategy, standardization and open-source domains. Alla has over 25 years of experience in developing technology strategy and system architecture at global companies, and holds a Master's in Technology Management from NYU, and a Bachelor's in Mathematics and Computer Science from Tel Aviv University.

Cesc has a PhD in computer architecture, and is currently Intel Principal Architect inside the NEX organization. He's driving creation and definition of current and next Edge end-to-end system architectures.

Welcome, Alla and Cesc, and thank you for taking the time to join us today. I will hand over to Alla to start off. Thank you.

Alla Goldner

Thank you very much, and it's my pleasure and an honor to present on Intel Network Builders' platform and I would like to thank Intel for providing such an opportunity.

Now, what to start from? This topic, I would actually start from recapping the key changes that are accompanying the rollout and deployment of 5G networks and enabling it to become a truly programmable network. So, we would see several different aspects, several different domains of it, of actually how to become a programmable network, or what is needed for programmable network. And the first and probably most significant one is the cloudification of the mobile network. It started in the core with traditional NFV. We have seen that also in 4G, with some network functions moving from physical to become individual, but now because 5G calls for true cloud, to embrace cloud-native technologies, it is really picking up, and the industry is entering a new phase of network cloudification. Now, in parallel, industry initiatives such as O-RAN and TIP's Open RAN are making significant steps in opening the RAN. A diverse ecosystem of vendors and organizations promote C-RAN and vRAN, which are stones on the way towards fully standardized O-RAN deployment, and those solutions facilitate disaggregation and cloudification of the RAN action. This altogether is causing network cloudification to move rapidly from the core out to the RAN and network Edge, and requiring service providers to invest in building these new network clouds.

Now, the second big trend is that 5G core itself, as standardized by 3GPP, which is service-based architecture, and is leveraging the cloud-native principle in the way it is being built as a service-based architecture, in order to enable a distribution of the network function and dynamic scaling of the network functions. Now, service providers are also looking to open up their 5G networks and deploy a network function from multiple vendors, and operators anticipate these changes will enable faster time to market, faster time to innovate, to deploy and deliver new services.

Now, the additional aspects, which is the key of the way, also, the 5G architecture is built end-to-end in RAN, in transport, and in core is actually end-to-end slicing, which enables to support diverse 5G services of a shared network infrastructure, and essential network slice is a logical network that serves a defined business need by means of a particular set of network capabilities and characteristics. Network slice, as I said, spans across RAN, Edge, transport, and core domains, and they utilize various network functions and resources within each domain.

Now, one additional major aspect is about monetization, and as in the previous generation, the way to monetize was mostly about the charging models between the subscribers and service providers. 5G brings a lot of the new monetization capabilities based on 5G standards, such as slicing, charging, such as Edge charging, where actually you're not just reporting this performance on subscriber level, but there is also a possibility, in case of slicing, for example, from the orchestration system, to report on consumption of the network or consumption of the slices, and that really opens the door for the new models of monetization between the service providers and enterprises, which are anticipated to be the major customers of 5G.

Nevertheless, many parts of the network will continue to rely on traditional physical appliance-based network elements in the coming years, along with those virtual and cloud network functions. As we know, there will be still some physical network content, and most importantly, there will be a blend of all this together-- physical, virtual, cloud native, the mix of those different network functions from the different types of different domains offer different types of the clouds, public/private clouds, public clouds provided by different cloud providers, which are managed by horizontal service operation centers. That's quite a complex task, and that's really the core of why those programmable networks are needed to manage this complexity. This complexity impacts service provider's ability to innovate and offer next-generation digital Edge services that meet their customers' emerging needs, and the way to manage this is to have programmable manageable systems end-to-end, which takes care of different types of cloud, of different types of the network functions, as I said, and manage it in a centralized manner.

With that said, I'm moving to the next section, which Cesc will provide.

Francesc Guim

Thank you, Alla. So, yes, I think that you covered a very interesting point, because when you look at the Edge, and with the appearance of 5G becoming more available and deployed everywhere, I think that one of the fundamental aspects that you talked about on the scalability, on the dynamicity, on the non elasticity of the... Edge is really a live entity, and I think that when you do architectures, like long time back, we used to look at the Edge as a single point of location, but with the incumbent of new deployment models that customers are looking at from device down to the cloud and data center, we really need to understand Edge as a continuum, and of compute that goes from the more left to the more right of the spectrum.

Now, it's very important at the end of the day, when you do Edge architectures, I think everything comes into question of the cost. On why I'm putting services in a particular location, it's because of some KPIs that are dictated by the use case, and Alla is going to talk about the use cases in a couple of slides, but really understanding why we are putting certain services in a particular location. It can be latency, it can be data privacy, it can be IP protection, it can data sovereignty. There are multiple KPIs that you can use to map the use case in the proper location, and in some cases, that mapping will be dynamic, and depending on the load that you have in that particular location, the load that you have into the network, and the loads that you may have even into the cloud, and to this point, it’s very important that when we do orchestration architectures, that we look at this multi-tier aspect and always trying to be aware that services may move from one place to the other, depending on the dynamic aspects of Edge.

I think that when you look at the KPI that we typically measure when we design those systems, it tends to be-- like TCO is a good one, because once you establish the KPIs that basically can hold-- let's call them more requirements than KPIs that dictate where the services can be executed, and why, now you have to do the orchestration and the management based on the TCO and, basically, with this dynamicity, try to be effectively using the devices and resources, and trying to minimize things like power, try to maximize service density, and try to maximize the service-level objective or QoS that those services are achieving. And that's why then all these things together, you need to combine the scalability of the system itself, but also the consistency from service perspective.

And something that I think is very critical for us as an industry to be aware is that in 90 or 95% of the cases, the users that are developing applications, they're going to be very sharp and very good into the domain of the application, but they don't need to be aware of what's happening underneath, and whether the application is going to be landing into a platform that has BPU or GPU to accelerate the media and AI, or whether the platform is just a CPU-based platform, and all these types of things, or even they don't need to be aware that they're using a GPU that is shared with other tenants of the system. So, that's where it's very important, when you look at all these elements, that you have a cloud native-- like the combination of a virtualized and cloud-native solution, depending on your deployment model, that really takes all these things into account, and always seeking the ease of deployment of the applications and the abstraction of the complexities.

And in this case, cloud native implies that the software stacks need to be aware of how to manage the virtualization of the discrete resources, it implies that the orchestration and stacks, they need to be aware on how to manage things like SR-IOV or virtual functions for the networking, and try to isolate traffic, or even how to manage the CPUs and do CPU core pinning on the services, depending on the quality of service that they require.

When you look at this problem, per se, it’s really complex. We're looking on devices to the data center, but luckily, as the ecosystem matures, and we have orchestration capabilities as the ones that we're presenting today with Alla, it's also good that the ecosystem is providing new constructs in the form of C&C microservices or things that we can use for managing the devices. For example, from Intel’s perspective, we provide operators to manage the lifecycle management of the networkers, plugins as well to do the management of the accelerators and networkers. So, there are many different elements that the ecosystem is developing to make this orchestration and infrastructure management more easy to deploy and to implement.

Now, another important aspect that we have to get into account is that not only Edge is dynamic, when you look into the loads that the systems will have, the loads that the network will have, or the different tiers that you will have in the Edge, each of those tiers having different types of hardware. Also, we have to be aware that the more that Edge evolves over the time, it’s not like that it's going to be one single type of Edge where you run one type of workload, like a telco Edge or IoT Edge or enterprise Edge. The things that we are seeing

more and more is that, at the end of the day, many of the customers that may be more focused on a particular vertical, they will always end up having use cases that can be mapped across different verticals. You can have an example of a car manufacturer customer that has the most of the use cases anchored around maybe the PRC management for the robots, and maybe some shop floor video analytics for the robots that are AMRs that are moving around, but maybe as well, the same manufacturer is willing to deploy 5G video analytics to manage the cars that get produced into the shop floor, and need to be parked autonomously into the parking lot.

And at the same time in this multi-service type of deployment models, we have to be aware that it's not only everything about one type of access. So, things that we see more and more is that there is a combination of different types of networking or data plane capabilities, that you may have some devices that are being connected into the Edge architecture, using like 5G cameras, for instance, that will have to go into a network, maybe FlexRAN plus a network function on the UPF, and then do the traffic streaming into the service that is processing those camera streams. But at the same time, you may have some sensors or some other cameras that are connected through traditional MPLS network or ethernet network and, in this case, the network management is different, but the service management is common.

So, it's important that we have this kind of, what we call, Common Frameworks, so that we understand how these commonalities in the deployment models on having different infrastructure, let's say, communication architectures and having a common service management architecture, it's very-- it's relevant, and it's a use case that is happening more and more. And that we see that the orchestration solutions are being more adaptive into this type of more heterogenous deployment model.

So, now, let me hand it over to you, Alla, again, and continue with the use cases.

Alla Goldner

—mentioned some of them require low latency, some of them require high bandwidth, really density, specific requirements. Those use cases can be in emergency response. By the way, speaking of emergency response, there are some famous use cases, which also have O-RAN and RIC support. Energy use cases, which I will also show some of them, then I talk about TMO Lab where we participate. Obviously, education use cases, smart city use case. In finance area, health area where, basically, some health applications which require ultra-low latency can be provided. All the industrial IoT specific use cases where each of the requirements, each of the applications would have its own requirements. Media. Retail. Smart homes, obviously, were actually different under smart home, different services may require different type of service and, basically, different types of the slicing. And some of them would require also Edge nearby to support those with the low latency.

Smart cities, of course, you know those-- we are going there. There are already deployments in some cities of those services which actually require those different latency, throughput, density support. And basically, each one of those would actually also require-- some of them would require Edge deployment, some of them would require, simply, slicing, and for some of them there will be specific slices where, actually, Edge would be an integral part of them.

Now, speaking of the capabilities of OSS and network management. As you remember, I started this presentation from actually describing what programmable networks are needed for. And summarizing that slide by saying that the end-to-end network management solution in the OSS place is needed in order to bring up that programmable network. But saying end-to-end is simple, now what it really means?

So, it composes of several parts, we believe. One first part is about service design. The second part is about orchestration. Really, it is pretty difficult to disconnect between those two, they go together. So, if you look into design and orchestration and what they provide, they actually bring the capability for services, which are designed with the help of design tool, and enabling by stitching together various relevant network functions and applications. So, basically, we design those services or we design slices, we design location of the network functions, which are needed to fulfill certain service requirements, and then we manage the lifecycle of those services and of those network functions by having the orchestration system.

Now, the additional key part of the solution is, of course, real-time active and federated inventory, where we have-- we provide a comprehensive close to real-time view of the many more moving pieces in the network, all the resources, all the network functions, all

the capabilities are here. And of course, and we see that coming, we see an increasing demand, I would say in the last year, and now all the discussions with service providers are pretty much around it in the Closed-loop Assurance system based on AI/ML where, basically, all the orchestration decisions are based on those inputs and recommendations. And eventually, mechanisms such as scaling, such as healing are done in advance, and not then the problem occurs, so problem does not occur. And that is really common. If two, three years ago, that was probably a marketing buzz to speak about assurance based on AI/ML, and only statistics-based assurance were basically in place or needed. It is not the case any longer. It is really advancing towards AI/ML system-based, not just on statistics, but on predictions and recommendations with the help of machine learning-based system.

Now, speaking of our solution, what we provide as Amdocs in this field. So, the key part of our solution is 5G slice management and Edge orchestration. As I said, this is in order to design, instantiate, and manage the lifecycle of the slices. Again, Edge location might become part of those slicing, and the Edge will also-- Edge support will also develop smart placement and homing system, which provides weights to several parameters, and based on the algorithm which calculates the overall recommendation. It also decides how to instantiate and how to scale those relevant network functions. I'm going to talk about it also a little bit more, shortly.

We provide inventory, of course, which is-- we have that inventory business for a long time and now, of course, it is going through transformation, it became a cloud native supporting 5G. The inventory is a single reference for network inventory covering all domains, RAN, transport, core, accurate view of hybrid services, Edge data center resources, basically, slices. Everything which is involved in the network and common data with visualization and federation.

And as I said, assurance. This is something that we also are looking on, not just as one single solution, but there are, I would say, assurances, a layered solution. Just as an example, NWDAF, which is a part of 5G core is in core. Non-real-time, near real-time, RIC, are a part of O-RAN architecture. Now, is the advancement of 3GPP and architecture, network management assurance is part of, actually-- of the network management layer and defined-- being defined by 3GPP as we speak. And actually, Amdocs has business in all those layers of assurance. We are building all those types of the solution. And also, the challenging piece is, actually, to connect those solutions and to see the consolidated picture.

As we build all of them, we actually believe that we are in a good position to actually interconnect all those solutions and to close the gaps between those solutions, which are not yet defined by standards.

Just an example, RIC, which I have mentioned before, and MDA], in RAN domain, what is their relation and how are they interconnected, how MDA is interconnected with NWDAF, how O-RAN works with NWDAF and so on. So, we are building all those solutions as a part of our assurance.

Now, a little bit of zoom into the slice management. So, we provide both slice management on the domain level and on end-to-end level, basically, interconnection, RAN, transport and core. We believe that as we are vendor-agnostic solution, we do have some network functions that we develop in-house, but we are famous for our OSS-based SaaS solutions as well. We are in a good position to actually bring that end-to-end management system for slicing, built by 3GPP standards. And here on the slide, you can see those 3GPP standards, 3GPP, ETSI NFV, TMF APIs that we are supporting in our product, which basically interconnects RAN, transport and core. We are already working with the key network equipment providers in order to build those solutions, several POCs that were already ongoing and, of course, participate in the different RFIs, RFPs.

Also, in addition to that horizontal integration, we provide vertical integration starting from BSS all the way down through OSS into virtualization. And this is because we believe that the close-coupled integration with the charging system, with the ordering system is the key also for a good network management solution.

For slicing, and one additional thing, as I've mentioned, you know there is a slicing and charging, there is an Edge charging defined by 3GPP, those are also parts of our network management solution, which basically reports to the charging function.

Now, for the Edge computing. I also mentioned on-- not the previous slide, the one before that. So, speaking of the Edge computing, I think that one of the key features is placement and homing, and that definition goes for actually enabling the algorithms for deciding

where to instantiate and where to scale and how to scale the different network functions and the different locations. And so, we build it by supporting an algorithm which supports parameters such as low latency, latency in general, different hardware accelerators which are needed for certain network functions, availability which is required, customer preferences of course, affinity, affinity for different network functions. And based on all those parameters, basically, these are the weights which are provided by the service provider depending on the needs, we calculate the best placement and scaling for the network function. And this is how we instantiate and scale and activate that.

Now, one example I wanted to share with you on the specific work that we are doing in TMO Lab, this is the real thing that we are doing in the real lab, this is the real vendor. So, the use case… TMO is T-Mobile, of course. We are working, also, with other partners. We are working closely with Azure there as well. So, Amdocs is one of the founders of this lab. The goal of the lab is to build a software ecosystem focused on the enormous applications, services, and use case potential for Edge and for 5G.

So, the use case that we are working on is an agriculture use case designed to provide farmers with modern data collection solutions and insight that will enable them to improve their practices, increase their yield, and cut costs. In many rural areas, growers are unable to adopt new technologies and practices, obviously, because they don't have appropriate connectivity. For this use case, we demonstrate how 5G and Edge computing can bring a tremendous amount of value in helping one of the largest industries in the world increase productivity and profitability.

So, Amdocs Orchestrator, in this case, is responsible for managing the lifecycle of the deployed application. Drones… so, that use case consists of drones, which fly over fields and orchards and take images, which they stream to an image analyser, an application from Microsoft running at the Edge.

In addition, we have IoT sensors deployed. We provide data on soil nutrition and humidity, chemical, and more parameters. The images and data collected by all these IoT sensors are processed at the Edge, and then the processed data is sent to another machine learning platform that runs on Azure Cloud. This AI-based data analytics application then takes all the data from those multiple sources and processes it to provide farmers with more accurate insight that actually empowers them to make better decisions concerning labor planning, irrigation, scheduling, pests, disease control and so on.

The solution also included AR-based remote equipment maintenance and safety security video analytics application that is also deployed at the Edge. There are many, many interesting things that we can speak about and describe about this lab, so I suggest that if you are interested, approach me but, perhaps, offline and I will be happy to provide you with that information.

And with that said, I am moving back to Cesc.

Francesc Guim

Thank you, Alla. So, I just… before getting into a little bit on the Intel technologies, I just wanted to reiterate some of the points of what Alla was talking about during the last few slides. I think I kind of already emphasized at the beginning the interplay between the networking and the compute, and I think all the technology ingredients and products that Alla was talking about that, basically connecting in an end-to-end piece and connect to the networking elements, the network slicing, the interplay between the services and the network, these are one of the fundamental KPIs and design choices for Edge. Because one of the major differences that you see when you get into deploy end-to-end Edge systems architecture with respect to traditional hyper-converged data centers where you are hosting everything into a data center and consumers or producers tend to be into the same data center here, really, the use cases dictate that you need to have a more fine grain control on the networking and how the bytes are moved from the producers that can be devices into the services themselves, or even from the consumers if there is… like the consumer is, for example, a gaming application that is being executed in the Edge, but consumed by an Edge device like mobile or any other type of devices.

So, here, the interplay of the telemetry and orchestration of the network becomes a very important aspect.

Now, connecting that with what's coming next, one of the aspects as well that goes into this end-to-end picture that we've been talking for 30 minutes now is that, obviously, Edge-- depending on the Edge location that you have, you have certain types of requirements and restrictions that basically go-- go basically like dictated by that Edge location. If you are more close into the device, typically, you will have more-- lower latencies and higher bandwidth. But obviously, the type of compute that you can typically place more on the far Edge is going to be more small form factors and more power and thermal limited. And as you move more close into the data center, you will have more traditional type of compute requirements, like maybe half a rack or full rack of traditional data centers maybe with certain capabilities more tied into Edge like, for example, certain NICs that provide better network management or things like that.

But at the end of the day, you need-- what it implies is that you will have different hops of the Edge that will have different requirements, and those requirements will be fulfilled by different types of compute and storage technologies that are optimized for that particular deployment model. So, if you go, for example, more on the far Edge where you need to have extended temperature support between minus 20 and maybe up to 70 degrees ambient, so that means that you will have to have CPUs, like Xeon D based, that they support this type of extended temperature, or chassis from OEMs that have this level of support.

And also, as well the type of accelerators that you will place on those deployments are going to be more specialized and more performance per watt efficient. And as you move more into less constrained deployments, you will have more high-end GPUs and high-end BPUs that can accelerate with less constrained thermals and power envelope. But obviously, with more overall consumption and more overall throughput. And that means that you will have to know the services that you want there.

It’s important that when you do your system architecture and you have services that are mapping to these different locations, you need to understand the scalability of these services in each of these locations. It does not necessarily mean that-- and I'm going to talk a little bit on OpenVINO and why OpenVINO is used to abstract at a certain level what some of these complexities of what this underlying hardware provides. But it’s important that when you have different types of Edge appliances where your services can run, that you understand how do they scale. So, when you do the orchestration and placement of those workloads, you understand what resources are required.

And as an example, some of the work that we do is now you have, for example, a video analytic use case that is doing surveillance, and you know that it has to do maybe 11 frames per second per stream. Now, if you map this use case into a platform that is based on a Xeon D, you may be able to maybe have compute maybe 10 cameras per-- 10 streams into this platform. But maybe now if you add a BPU on top of this platform, maybe you go up to 40 streams per-- in terms of density. And if you move now into maybe-- deeper into... more closely into the data center Edge, maybe you can go up to 80 streams per platform using a standard-- more high-end Xeon.

But these are the types of things that are important to understand. And this is... we have a variety of reference implementations and elements within Intel, like AppCloud and things like that, that allow the ISVs to understand the scalability, but as well, allow to get this level of characterization that can be used from an orchestration perspective.

And when you look at the... Intel is not only about compute. So, we look at the end-to-end systems and looking from the networking aspect and looking into connectivity from network cards, silicon photonics or switches like Tofino switches. We look at memories and storage and how to make data persistent and durable and, basically, looking at tradeoffs of, OK, how my workloads are utilizing the storage, the memory, and how do I define my platform based on those requirements.

And then, obviously, we go into other areas where we have a Xeon, Atom and Core, which are different CPUs or processors that have different capabilities. And depending on your deployment model and your requirements, you will choose one or the other. And along with those things, you will have different accelerators that provide different functionalities.

Something important here is that that may look like a cumbersome type of puzzle, because I'm talking about, OK, I have 12 different options that can be combined between themselves in multiple ways. So, as part of the Intel work, so we aim to facilitate the adoption for those technologies in a more easy way. And there are a couple of areas which are important where we’re working on.

One, it's on-- first of all, on the abstraction on hardware accelerators and the compute, the CPUs are utilized by the services themselves. And in this case, you can look at and think like oneAPI, that basically abstracts a little bit the complexity of the programmability of those devices using common APIs. And as well, frameworks like OpenVINO.

So, OpenVINO, for the ones that you guys are not familiar is a framework that allows you to-- you develop your video analytic application that can be used on top of things like GStreamer or DLStreamer that are standard ways of implementing AI pipelines. And now, you can-- with the OpenVINO framework, you can access into different types of accelerators and hardware capabilities in a more or less transparent way.

And what we are aiming here is to facilitate that developers have one single implementation, and this single implementation can land into different locations of the Edge. And then now you have elements of the framework that transparently will use and take benefit of the hardware acceleration capabilities underneath without the need of modifying the pipelines and the applications themselves.

And some examples that you can look up, it's things like Video Analytics Serving, there are microservice-based implementations to, basically, have access into these compute elements and where you can reduce your AI pipelines, media and AI pipelines and make this in a transparent way.

Now, the second aspect as well is not so much into the developer from application perspective, but on the developer on the infrastructure area. And here is where we have elements like technologies such as smart Edge where we are developing plugins and operators and cloud native technologies that, basically, are meant to make easy to deploy and manage all these different hardware technologies that we're defining in the context of Intel and Edge technologies. A lot of the engagements and joint work that we do with Amdocs to, basically, upstream these elements from the service perspective, and infrastructure perspective in something that can be deployed and that has productization pathway.

Just a little bit on-- just to provide some more view on what I was talking about, the OpenVINO toolkit and things like Video Analytics Serving. If you look, there is some reference in the last slides of the presentation but, basically, this is one of the examples along with oneAPI that are... we provide this easy to deploy framework that, basically, you define your own AI and media application, and now you have different optimizations for different hardware and now you can deploy those things in a more transparent way.

And I think that if you look at Edge perspective, like 80% of the-- I would say, it's a number that I'm just making it up a little bit, but I would say that 80% of the use cases that you will see are video analytics-based. There's a lot of content delivery, for instance, and other use cases like augmented reality. But video analytics is a big chunk of the use cases, so you can get a little bit deep into how to utilize OpenVINO to make more scalable and portable applications that can be orchestrated in the fashion that we have been discussing so far.

As well, like when you look in that-- a little bit on getting some more details on the OpenVINO, you don't have to think of this as something that is just... I get to implement my video analytics application. Also, we provide mechanisms to train models and to optimize models and to get the best density performance, density of the platforms that we provide. It's something that-- it's very important when you look at the connector, the DCO that we were talking at the beginning, the lifecycle management of the services is that these applications, you can easily get 20, 40% more performance improvement if you are using the right optimizations at the hardware level.

And all of those things, obviously, as I said a couple of slides earlier, it doesn't imply that you have to have deep technical understanding on the hardware, but really understand what techniques you have to utilize and what tools and what applications.

In the context of OpenVINO, you can find things that allows you, for example, to quantify models. So, for the ones that you are familiar with, with AI models, you can have neural networks that are working with FP32 resolutions, but you can quantize them into INT8, and if there is no change into accuracy, you can get performance improvements up to 40 or 70%.

And basically, if you go a little bit deeper with the OpenVINO toolkit, you will see that we provide all these end-to-end capabilities that go from building models, optimizing them, and deploying them into different types of hardwares and Edge appliances.

Now, it's also interesting that when you look at-- as I was saying on the infrastructure side that you understand that it's not only about the services that we have to have inside the infrastructure and service orchestration, all the different elements, basically, allow those services to be deployed and allow those services to access in a transparent way into the hardware capabilities. And when you look at the Smart Edge Open that previously was called OpenNESS, you have access into multiple cloud native constructs that provide access into Kubernetes operators and plugins that, basically, allowed for the infrastructure and service orchestration to manage all the technologies that I just mentioned. And also, we provide access into elements that provide global observability for the telemetry and monitoring, so this type of lifecycle management of the application is satisfied.

And also, always looking at multi-cloud, so it's from the access to the cloud and connected into the examples that Alla was talking about, it's really on orchestration in a multi-tiered domain.

So, let me hand it over to you, Alla, again and talk a little bit on the journey between Amdocs and Intel in the Smart Edge Open.

Alla Goldner

Sure, my pleasure. So, now we will kind of summarize it and describe what we are really doing together. Already Cesc has mentioned, of course, but this is in a more formal manner what it is, indeed, our solution provides.

So, our solution of the slice and orchestration is powered by Intel Smart Edge Open. We joined the forces and our partnership goes in many different forms, but... *[technical difficulty]*... of course is very important.

So, we create a solution that brings the benefit of cloud native operation model to the Edge based on our automation platform, which includes, as I explained, orchestration, assurance, inventory and design. We enable distributed 5G core deployment, which uses Intel Smart Edge Open features, and allows intelligent placements and homing of those network functions at the Edge per requirements of those applications, and requirements from the service provider for which we deploy that. Delivering advanced services, as you know, also was mentioned in the example that I provided for agriculture, real-time video surveillance, and all the range of other applications, enhanced with hardware optimization provided by Intel Smart Edge Open at the Edge of the network. And Cesc talked about it and described what that means.

Of course, the idea here is to provide an integrated solution, which brings a reduced time to market and lowers the investments of new Edge services deployment through that unified architecture, which I am going to show shortly.

And what we target, really, so we target a fast onboarding and deployment of network functions, services, and application that are enhanced by Intel Smart Edge Open, basically, by pre-integrating some of those solutions and bringing that solution already in place into the market. Unified architecture, of course, relying on Kubernetes with Intel Smart Edge Open support.

Speaking of the network functions and network equipment providers, cloud partners, we already work with some of them and that pre-integrated solution includes already support for some of them. We can, of course, extend it and we would be happy to extend it based on the specific service provider's demands.

Now, that's the same, but showing that picture of our integration as a joint architecture. So, you can see-- on the right side, you can see our Slice Manager platform supporting end-to-end through RAN/Edge, transport, core towards the internet, managing the lifecycle of slicing end-to-end with some based on Edge deployments.

On the left side, you can see those tools and accelerators and controllers, telemetry tools, all those things that Intel brings as a part of their solution. So, we're combining those two pieces. We integrate our solution with the platform provided by Intel in order to accelerate time to market and Edge deployment at scale, and that's really the idea of the solution.

And you know, just as the last thing that I wanted to show here is that, basically, the example of how that consolidated system of Amdocs' solution with Intel's solution works.

So, in the example that I show now, you can see a deployment of a 5G network, which relies on multiple NearEdge nodes. Those nodes are being built by actually Intel open Edge architecture. In addition, there is a central data center where the signal and core functions of 5G core are deployed. Also, there is a data center where the Amdocs Orchestration and Management System is deployed.

Now, the Amdocs NEO, which is a Management and Orchestration System handles the creation and management of 5G network slicing and 5G network slice subnets in each of the domains to satisfy requirements of some communication service or application by other words. Example scenarios for such a service based on 5G network slice can be event, venue slice, factory slice, enterprise slice, as we said, smart city, IoT-- different IoT application, agriculture application, pretty much all of the services and applications. Amdocs orchestration chooses on which near Edge node to place latency or bandwidth sensitive network functions based on rules, policies, machine learning outcomes, and need of those functions including and, of course, very important, platform acceleration features that, again, Cesc was talking about and provided by Intel as a part of their solution. Amdocs new orchestrator creates and manages the lifecycle end-to-end lifecycle of 5G network providing such network slices, again, based on those features and tools that Intel provides.

With that said, I would like to finish our presentation and move to the questions.

Lilian Veras

Excellent. Thank you, Alla, and Cesc for sharing such great information with us. We do have a few questions coming in from the audience, so let's get started on those.

The first question I have here, a member from the audience asks, "Are there any use cases of services that Amdocs service and network automation solution supports?"

Alla Goldner

Yes, so Amdocs and partners integrate Intel Smart Edge Open, Kubernetes operators to accelerate time to market and Edge deployment at scale by using hardware accelerator, FPGA, QAT, SR-IOV and others that Cesc has mentioned. We do have examples of those use cases on the slide. You can look on the presentation that I presented for more details and, of course, you can connect offline, you can connect me and, of course, you can connect Cesc. We would be happy to provide some specifics on the use cases which are of your interest.

Francesc Guim

The other thing that you guys can have a look, so as part of the engagement on smart Edge, there is the application hub that, basically, we work with the ecosystem players that they develop specific solutions that are onboarded into the Smart Edge Open. They, basically, are already onboarded into smart Edge. So, we have a set of partners for different domains from content delivery, to retail, to smart cities and other types of the verticals that they are part of the smart Edge family and that they can be integrated in any engagement with Amdocs and ourselves for different use cases.

The other thing that you guys can have a look too is there is some-- there is the Edge Software Hub where there are some use cases that they can be used as a baseline reference implementation for developing use cases for different verticals that you can have a look as well.

Lilian Veras

Great, thanks to both of you. Another question we have here, "How could we learn more on the capabilities of Amdocs' service and network automation solution?" Alla, I guess, this is for you.

Alla Goldner

Yes, thank you. I guess the best way would be to leave your contact details at the Q&A chat board, and we would contact you in the back and discuss your specific requirements and also more in-depth at the specific capabilities of our service and network automation solution.

Lilian Veras

Awesome. Another question here. “Are there any prerequisites on our network Edge infrastructure in order to Amdocs service and network automation solution, together with Intel Smart Edge Open, to be deployed?”

Alla Goldner

Well, as I mentioned during the presentation, we already have an integration with some network equipment providers, some cloud providers, some OSVs, which are a part of our joint solution with Intel. There may be, of course, additional demands for integration with additional vendors, and we would be more than happy to discuss it with you. And of course, the solution is fully open to integrate with additional partners, whether we are talking about network equipment providers, cloud providers. So, whichever customer, basically, customer demand is, we are happy to discuss it.

Lilian Veras

Awesome, thank you. We do have time for one last question. A member from the audience is asking, “What is the typical engagement model with communication service providers, partners for the planning, design, implementation, operations, operations and support, and training?”

Alla Goldner

Well, so you know there are RFIs/RFPs, which are coming from the service providers. Basically, in some cases, the whole network management and orchestration system is required-- provided by the single vendors. In some other cases, there are several vendors selected and I can tell you, there might be some for design, some other for orchestration or assurances. I can tell you that we participate in engagements of both kinds, basically. And again, if you have any specific question asking about some specific-- in order to understand better that engagement, I encourage you to leave your details in the Q&A chat screen and let’s discuss it in detail.

Francesc Guim

I know there are just 30 seconds as well to add on top of what Alla said. There are other types of deployment models that are important for us to keep in mind, which are on-premise, so like private 5G or in the US, CBRS. So, there are places where there is private spectrum that can be used for enterprises and end customers to deploy their own 5Gtest bench.

And in these cases, it’s really like a dialogue between the end customer and the one providing the solution, and providing the internal architecture. And in these cases, it’s really an engagement between-- in this case, it could be Amdocs and the end customers, and this is something that it’s interesting in terms of the deployment model and architecture, because it includes all the different elements that we discussed during the presentation today.

Lilian Veras

Awesome. Thanks, Alla and Cesc, again, for the great presentation today. Unfortunately, we are running out of time for the presentation. I would like to ask our audience to please not forget to give our team a rating for the live recording, so we may continuously improve the quality of our webinars.

This concludes our webcast. Thanks all.

Alla Goldner

Thank you very much.

Francesc Guim

Thank you, bye-bye.

Alla Goldner

Bye-bye.