Intel

How ASUS, Intel, and Silicom Position Edge Products Into 5G Architecture

CORPORATE PARTICIPANTS

Lilian Veras
Moderator

Alber Wu
ASUS – Server Business Unit, Division Director

Oren Benisty
Silicom – EVP Strategic Sales

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: How ASUS, Intel, and Silicom Position Edge Products into 5G Architecture.

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'll also find an Attachments tab with additional documentation and reference materials. 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, 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.

Today, we are pleased to welcome Alber Wu from ASUS and Oren Benisty from Silicom.

Alber Wu is the current Division Director of the ASUS Server and Workstation Business Unit. Alber joined ASUS in the year 2000 and has built an extensive and enviable experience of developing server products for the telecoms industry, from the initial design stage to the production of world-leading complete solutions.

Oren Benisty is Executive Vice President for Strategic Sales at Silicom. Oren has a bachelor's degree in electrical engineering, and a master's degree in business administration. Oren has over 20 years' experience in the telecom and data network infrastructure design.

Welcome Alber and Oren, and thank you for taking the time to join us today. Alber, over to you.

Alber Wu

OK, thanks both Lillian's and Intel's arrangement on that.

Hi, this is Alber from ASUS Tech. I think it is my great honor to join the webinar. Today, I would like to share with you about what we are planning for the three-way collaboration and solution on 5G edge.

First, talking about the 5G opportunities. From the infrastructure point of view, no matter what kind of definition from the cloud to network, network edge, or on-premise edge, most important is how to leverage the 5G features as the low power, low latency, and the fast speed with more connection. Then apply such solution on a vertical like industrial, medical, and the traffic that we enhance the convenience for live implement.

However, due to the 5G edge, which owns the key features like open and flexible for development, and it covers a wide field like tower design, networking, and system architecture. So, that's why we initiated the three-way collaboration for the 5G edge between ASUS, Intel, and Silicom. We plan to have the ultimate edge solutions, which integrates with Intel FlexRAN and Silicom co-technology on networking. Then ASUS would consolidate all the technologies into the best fit for the 5G edge application.

This cooperation delivers better solutions with strict test teams and organization in hardware and firmware design. Later on, I will show you in detail.

Here, let's talk more about the joint design concept. As mentioned, all the designs is based on Intel FlexRAN architecture. Surely, we adopt the Intel network-optimized processors and networking adapters, then leverage the Silicom to manage the co-technology. You can see from the screen.

The first one is the Time Sync Technology to ensure the accuracy and the synchronization while data transfers. And the second one is FEC Technology, which enhances the quality and the efficiency within transfer. Besides that, ASUS focuses on the design of the DU and CU servers to consolidate all the technology into available and flexible solutions, which contain the key features like short chassis from ASUS and the NEBS compliance.

Along with ASUS, the hardware design experience and always the co-technology. ASUS edge servers deliver the ultimate design to the telco market.

OK, here, as the previous introduction, our design is based on Intel FlexRAN. This innovative architecture provides the prosperity on Intel-based server to penetrate the telco market. It allows us to easily integrate on top of the software architecture, and the superior Intel processor performance.

On the other hand, ASUS also engaged closely on the Intel Select Solutions certification. This can also shorten gaps while implementing the FlexRAN product to the telco market.

About Silicom, I think that they deeply focus on networking applications, like Time Sync, Forward Error Correction/FEC, and the SmartNIC technology. These ingredients are pivotal to the 5G application in monitoring, management, and enablement. So, ASUS feels so grateful to have a partner like Silicom. We try to have the testing and embed it on each of the products. No matter on ASUS general purpose servers or the proprietary ones, Silicom product works very well on it. Our partnership is so great to trigger adopting in the 5G segments.

Here, let me introduce you about the ASUS Core Competency. ASUS servers is not just rooted on the hardware design, but also they're pursuing greater performance and economic power consumption through the dedicated team efforts to create that exclusive tuning technology and innovation. That's why ASUS Server can get over 850 for the worldwide number one benchmarks records in the SPEC CPU.

Meanwhile, ASUS didn't compromise in power consumption for the green purpose. All these technologies are easy to access through the BIOS or the BMC.

As mentioned, ASUS product portfolio can meet different use cases from the cloud to edge. You can see here on the right side, this high density server for high performance computing. Also, we have the AI GPU servers to expand application in the AI channel and inferencing. Along with the common 5G here, ASUS also developed edge servers like 1U short-depth and comprised multi-nodes as well. More complete portfolio would be coming soon in the near future. Let's have some overview on the ASUS server and edge solution in the following pages.

Here you can see that ASUS servers always features on the latest technology and implement into the hardware and the firmware design. We would like to keep all the products have a focused strategy on the flexibility, scalability, and versatility. Among all of that, you can see here is the balanced design. This architecture is one of the key features which allows the end users every penny counts. That means

ASUS fully utilizes layout to maximize the scenario coverage and gives the best advantage to customers to their applications. On the other hand, we also support releasing our essential technology through the BMC. Then you could get the setting on your preference.

Last but not least, you could get hardware Root of Trust option in our 3rd Generation Intel Xeon platform too.

Then within edge server design, ASUS also offers the 1U short-depth, you can see from the right side here. And multi-node edge as well is the left side, we call the EG3050. To comply with the critical challenges in ambient temperature and location, which owns unified features like front access for the convenient maintenance, and DC/AC-IN to meet the different location requirements. Most important of all, passing NEBS to ensure the edge servers satisfy the regulation in telco industry.

Here, you can see what we provide and adjust on the hardware feature. It's also important on the high availability of the firmware for your application. So, we dedicated to address you all the available firmware in three ways.

The first one, easy to maintain. ASUS 5G server design completes a failsafe mechanism in firmware including BIOS and BMC, the dual images to low impact in sudden shutdown, and reduced overall TCO.

The second one, easy to safeguard. We also deploy the hardware Root of Trust to block malicious firmware attacks. Security is the top priority to ASUS servers, it is always.

And the third one, easy to deploy. The ASUS can work with independent software providers to ensure the full compatibility between the hardware and the software. In fact, ASUS, all the servers have to pass tests in a set location from the key platform providers like Red Hat Linux for Realtime. Also, with Wind River Cloud Platform. These platforms are standardized to meet the telco use case and worry-free to launch a new service.

OK, that's my section. Then I will hand over to Silicom, Oren. Oren, you go.

Oren Benisty

Slide number 12. Great to have you all over here today.

So, when we are looking at the 5G market, it's great for Silicom to collaborate with ASUS. ASUS provides the servers. As you can see in those pictures, we can see the server already equipped with dedicated add-in cards as accelerators. You can see the branch – broad range of add-in cards from standard eASIC cards based on Intel ACC100. We have FPGA cards for accelerator of Layer 1, and we have Time Sync. In the next slide, we'll get into more details regarding all these cards.

On top of that, Silicom also offers an IPSEC card for encryption. We've seen a couple of operators looking to encrypt all the traffic from the DU to the CU, and from the CU to the cloud.

The last card that we'll present today is UPF accelerator for the 5G core that they can accelerate up to 400-gig of traffic.

When we are looking at the 5G market, we are seeing all the networks connected using standard off-the-shelf servers. We are seeing standard add-in cards PCI Express, and all of these serve the market with Time Sync connectivity. So, we are seeing a GNSS receiver on top, the GPS satellite provides the clock into the DU. The DU can provide this clock into the 5G radios.

On the left hand side, we see the radio units. Those can be connected with eCPRI based on the O-RAN standard and today, we are in development of FPGA cards for the open fronthaul.

Moving to the next slide, we get into more details into the card. First of all, you can see a full server equipped with a Time Sync card on the right hand side, with 1PPS in/1PPS out, 10 megahertz in/10 megahertz out, and we have time of day. So, the first card that we'll discuss is STS.

STS is Silicom Time Sync. Silicom Time Sync provides connectivity based on Intel E810. E810 is ethernet controller, code-named Columbiaville. It supports both 1588 and SyncE. The line of STS cards includes a range of cards with four, eight, and 12 ports. We have a GNSS receiver on most of our cards, so for cost reduction, we have some cards without GNSS, but most of our cards include a GNSS

receiver. It means that we can be Grand Master. We support both Boundary and Transparent mode, so we can function to recover the clock both from the fronthaul or the backhaul. We have hold over of four to eight hours and in a specific temperature, we are able also to do 12 hours and even 16 hours. Again, it depends on the OCXO that we have on the board.

The BMCA is a very interesting feature. BMCA is Best Master Clock Algorithm. When we are discussing clocking, we can get clock either from the backhaul, it means from the transport network. Some of the radio heads also include the GNSS receiver, so we can recover the clock from the fronthaul. We can have a GNSS receiver on the card itself. And all of this creates a range of clock sources.

Silicom implements this Master Clock Algorithm in order for the card or specifically the DU to provide the best clock in the system. This means that we are able to recover the clock from any source based on its quality. Of course, when everything fails, we are moving to the OCXO and we are doing hold over.

The second card called Pomona Lake, and the Silicom code-name, it's called Lisbon if you look in our website. And this one is based on Intel ACC100. ACC100 is an eASIC. It's doing both LDPC and Turbo. LDPC for 5G and Turbo for 4G. It integrates into the server using the BBDEV driver that integrates into the... sorry for moving to the next slide. It supports the BBDEV for the integration to the FlexRAN. We have a few versions of this card, standard temperature zero to 45, and extended temperature that ranges from minus five to 55. And of course, we have a BMC on the card to manage all the thermals on the card itself. This card is working in a lookaside model and I will elaborate in a second.

The last card that we are seeing on the left hand side is an FPGA card. The FPGA card, we have two versions, one based on Arria 10, and the second version based on Agilex. This FPGA card can perform a Layer 1 acceleration. It means that it can do LDPC, very similar to the Pomona Lake. It can do fronthaul connectivity with eight ports of 10 or 25, or 2x100. We support, of course, Time Sync, it has 1588 and SyncE. We have hold over, very similar to the Silicom Time Sync card. GNSS receiver, so we can be Grand Master. And we support the native clocking mechanism Boundary and Transparent modes. Of course, we have the code for the Best Master Clock Algorithm, so we can select the best clock on the FPGA card. And we are compliant to the O-RAN C1, C2, and C3.

When we are looking at all these cards, we have to remember one thing. In the 5G market, we have different deployment models. Some companies deploy it as a software-only, so they will use the STS card just for connectivity and Time Sync to the radio heads. Some companies are looking to do acceleration to the FEC. It means that they will use the ACC100 for offloading for the error correction. And this will be in a lookaside model. It means that the traffic coming to the system through the STS card moving to the ACC100, back to the host system, and then moving out to the CU.

The last card, the FPGA can perform functionalities way beyond FEC. So, it can do, first of all, the functionalities of the STS. It means connectivity to the radio head, including Time Sync. It can do forward error correction. And in the future, it will be able to do more functionalities as part of the Layer 1.

So, this is a range of platforms, a range of cards. We are able to accommodate any deployment model that the software vendor will decide. And more of that, we are able to provide software vendors the ability to differentiate, whether it's a software, whether it's a partial offload, or it's almost a full load – offload of the Layer 1 acceleration.

So, this is a very interesting range of products. All of these are integrated into the ASUS platforms and we are able to provide products for POCs. You can see the example right here in the middle.

Moving to the next slide, I'm always getting questions about the Time Sync capabilities. First of all, Silicom is fully compliant to the O-RAN requirements. We are following both O-RAN, TIP, and ITU-T. We are implementing on the Time Sync card, both the STS and the FPGA card 1588/PTP, both for IP Version 4 and Version 6. We support SyncE based on ITU-T 8262. You should ask yourself why we are doing both SyncE and PTP, and it means that we are able to do synchronization based on time, frequency, and phase. The system is synchronized based on PTP, and we are using SyncE to improve our accuracy.

The card, as I said, supports both Boundary and Transparent based on the 8273.2. And of course, we are supporting Grand Master based on 8273.1. We are supporting both PRTC and ePRTC, so we can do Class B for PRTC and we have an OCXO Stratum 3E. 3E is the

enhanced one, it means that we are able to hold over the clock, as I said before, in normal temperatures. We are able to do 16 hours, but we are committing for 12 hours, and in extreme temperatures, we are going down to eight hours. So, everything is dependent on the working temperature of the DU.

Of course, everything is coming with software drivers for the x86. We support Red Hat, VMware, and of course, Wind River. All those will have in their next release our drivers, and there will be seamless integration into the software layers.

Moving to the next slide, I will get now to the next level of details. The first card that I would like to discuss called STS1. STS1 stands for Silicom Time Sync #1. This card is the entry level card. It supports four ports of 10 or 25-gig. It's equipped with Intel E810, which is Columbiaville. Temperature range between zero to 45, and this card is half-height/half-length. Over here in the picture, you can see it's a full height, because we are putting the 1PPS in and 10 megahertz in as part of the bracket, but we have also a smaller bracket for installing in different locations. This one supports PTP and SyncE. Of course, Transparent and Boundary clock. And we have Grand Master capability. The Grand Master capability is done using external GNSS receiver. We are not equipped to have a GNSS receiver on the card itself because of its size. This card also implementing the Best Master Clock Algorithm, so we can have either clock from Transparent or Boundary, fronthaul or backhaul, as well as the 1PPS and 10 megahertz.

Moving to the next card, it's called STS2. STS2 is eight ports by 10-gig. This card, having all the features of STS1, with double the amount of ports. It means that if in the STS1 we had four ports, over here, we have eight ports. Eight ports of 10-gig. This one also has PTP and SyncE. And on top of that, we've got the GNSS receiver on-board. It means that we just need to connect to a GNSS antenna and by that we have the Grand Master capability within the card. Of course, we have 1PPS in/1PPS out, 10 megahertz in/10 megahertz out, so we can either cascade the system. We can provide external clock to the card, and we are able to do also the Boundary and Transparent, so it means that we have all the clocking features available. This card has the eight ports, so you can see in the picture, it's four ports of 10-gig, discrete ports. And we have one port of QSFP with another four ports. We have done this one mainly because of physical limitation of the panel.

Moving to the next one, it's the eASIC.

Just last comment about the STS card, so I'll move back to the previous slide. We have STS3 coming in the future. It will have eight ports of 25-gig, very similar features, and we have STS4 with 12 ports. The 12 ports have a very unique capability, because it's enabled us to remove the switch between the DU and the radio units. So, together with ASUS, we are able to build a server that has 12 ports connectivity, fronthaul offloading, and by that we can provide a system that does not require any external devices like switches, for example, like external clock sources. And we can provide a system fully ready for POC.

Moving to the next card, it's the eASIC. eASIC code-named at Intel called Pomona Lake. This card is based on the eASIC technology. It means that Intel put all the technology for LDPC and Turbo within an ASIC. And by that, we can create an offload engine into the KD and to the forward error correction. The Mount Bryce itself is fully integrated into FlexRAN.

The first version that supported the Mount Bryce, which is the chip of the eASIC was a version 20.08, as you can see in the slide. Since then, there is 20.11 and many newer versions, but this is the first one. So, if you are a software version – a software vendor looking to implement this one, make sure that you have the right FlexRAN software. Those can be downloaded from the Intel website. If you cannot find it, please send me an email and I will provide you a link to the Intel site, so you can download the right software. The card is PCI Express Gen 3 x16. Of course, half-height/half-length. We have brackets for full height and half height. We implement the BMC for monitoring and control, so we are able to control the thermals on the card itself using the whole system at BMC. We have a version for OCP as well. And the temperature range for the industrial version is minus 20 to 55. It means that some of the outdoor systems can use this card as well. It's optimized for 40 watts, so no requirement for external power source.

Moving to the next slide, it's the FPGA card. As you can see, the FPGA card is a bit more complex card. In this card, it's very similar to the Intel N3000. N3000 for Intel included two Fortville MAC and Arria 10. And what Silicom added to this card, we added the Time

Sync. So, we are able to convert old code that was running on the N3000 and also to provide the Time Sync capability. This card is based on the Arria 10 from Intel.

The next generation that we will launch in a few months will include an Agilex. Agilex is the next generation FPGA from Intel, and Intel already announced such a card that it will be in the future as part of its Technology Day. This card, today, is based on PCIe Gen 3 x16, supports of course 1588 and SyncE, and we have a GNSS receiver on the board, so we can be Grand Master as well. Because this card is optimized for power in a range of about 75 watts, so we don't need an external power source to the card itself.

So, all in all, what we can see that we provide a range of cards that can support any deployment model. It can support also capability to create both connectivity and acceleration for the Layer 1.

Thank you very much for listening and now we'll start answering questions.

Lilian Veras

Fabulous. Thank you, Oren, and Alber for this great presentation. We do have several questions that have come in while you were presenting, so let's get started on the Q&A.

First question I have here, "Does ASUS have any POC in telco segment?"

Alber Wu

OK, thanks, Lillian. This is Alber, sorry. Recently, I think that ASUS engaged with the Taiwan telco service providers to have the POC on the 5G private network, so that we would have more use cases and end users using it. And we will hope to work with the top vendors to create a real business case in the 5G application in the near future. Thanks for that.

Lilian Veras

Thank you. Thank you, Alber, awesome.

Second question. I have another question for you, Alber. "What are ASUS hardware design concepts for O-RAN servers?"

Alber Wu

OK, thanks. I have our senior leader, Cherrie here for the hardware architecture planning. So, Cherrie, would you help to answer this question.

Cherrie

Of course. Hello everyone. This is Cherrie speaking. ASUS core competency is that we design our own servers. We have proven performance tuning capability, excellent thermal and computing design experience. Hence, we also – we always optimize our products for different workloads. For DU or CU servers, we want to design products which can fit for the base station environment. So, the design question will be how to let servers run smoothly with high reliability and availability in a tiny space.

So, we designed our server with short-depth chassis and front access design, so a user can maintain at front side very easy. Also, redundant power supply with support to both DC-IN and AC-IN. Our server can be operating in critical environments and comply with NEBS regulation. Thank you.

Lilian Veras

That's awesome, thank you for that contribution.

Another question that we have here. "What design features – what are the design features in FW architecture?"

Alber Wu

OK, about this question, I will let our software leader, Joseph to have this answer.

Joseph

Hello everyone. And for 5G server design, so based on the use case and telco segment, we would like to have latest firmware, and normally the BIOS or BMC as a redundant, as a backup once the server is interrupted or in downtime, and the backup firmware could resume quickly, and to lower down the overall downtime and also can have the 5G service went back. And also on the other hand, we will have the BMC support RedFish protocol, which is much more convenient, also safety in overall out-of-band management. Also, for the PFR, and this is the great option hardware Root of Trust devices, which is able to help users to prevent malicious firmware attacks to keep the server in safe status.

Most important of all, our server will comply with the very important software vendors like RedHat and Wind River, and to be certified by their programs, and we can let our users to be easy to deploy ASUS servers in their fields. So, this is our design concept. Thank you.

Lilian Veras

Thank you. Very detailed explanation.

We have a question for Oren now. Oren, what is Silicom's key value proposing for the Time Sync cards?

Oren Benisty

OK, good question. Of course, there are a few solutions in the market. So, Silicom offers, first of all, software integration. Silicom offers all the drivers required for the Time Sync capability, initially integrating into the Intel FlexRAN, and after that, into major vRAN vendors.

We are also offering port density. We are offering four ports of 25, eight ports of 10, eight ports of 25, and 12 ports of 10/25. It means we have a range of products that...

Lilian Veras

Oren? Your voice was cut. Can you please repeat the end of your explanation?

Oren Benisty

Yes. So, I was saying that Silicom have a range of port densities as part of our cards. So, when a software vendor needs to validate a card, we bring a range of products with identical software. It means we have four ports of 25, eight ports of 10, eight ports of 25, and 12 ports of 10/25. It means that the software vendor has capability for every deployment model, whether it's rural area, urban, or dense urban deployments. All the same cards, all the same technology, all the same drivers. So, very simple to on-board.

The last thing that differentiates Silicom is the accuracy. We provide today the highest accuracy in the market. We provide the longest hold over time. Most of the solutions in the market today provide between four, and some of them also eight hours. We are able to provide even 12 hours and 16 hours. It depends on the card model.

Just to summarize, we have the highest accuracy, we have software capabilities to support the integration into the vRAN application, and we have port density to admit every deployment model. And of course, as I said before, we are able to remove the switch between the DU and the radio units, and provide all the capability through the card itself.

Lilian Veras

Perfect. Thank you, Oren.

Another question here for you is, "Why do we need both SyncE and PTP for time synchronization?"

Oren Benisty

OK, good question. I'm getting this one many times. So, this question comes up with two different capabilities.

So, first of all, the PTP or the 1588 is Layer 2/Layer 3. The drivers are running on the host who are terminating all the time synchronization in the host itself. In order to improve the accuracy of the synchronization, we are implementing SyncE. SyncE is implemented on the Layer 1. It means that it's in the physical layer, on the ethernet connection between the RU and the DU, and we are able to provide higher accuracy, that's why we are using PTP assisted with SyncE. So, both of these require for the accuracy that we are getting through these cards.

Lilian Veras

That's awesome. Thank you, Oren.

I do have one more question for Alber. "What features in ASUS 5G firmware architecture?" What are the features?

Alber Wu

OK, sorry, I will let Joseph to help answer this question on it.

Joseph

OK, so in our firmware design, when we were having redundance, so we would do some failsafe mechanism to keep the server state as in safe, and also quickly resume. So, this is the basic design. The second one would be the security enhancements. So, no matter in BMC firmware or in the PFR, so with these two kinds of solutions, we would like to improve the overall hardware management security. So, our key design here, all in all, it would be the redundant and also security enhancement in our firmware management. Thank you.

Lilian Veras

Thank you. And we have time for one last question. This question is here for Oren now.

Oren, Silicom eASIC and FPGA cards offer similar functionalities. What is the value for end customers?

Oren Benisty

OK, so again, if I go back to the slide, I'm going to move back to the slide. When we were looking at the different add-in cards for 5G, we have different functionalities. The eASIC provides best cost effective solution. The eASIC implements forward error correction for LDPC and Turbo, LDPC in 5G and Turbo in the 4G. This is a specific function that is offloaded from the host. This card is cost-optimized, power-optimized, and it provides the best value for this functionality.

When we are looking at FPGA, this is open architecture. It's an open... it means that we are able to implement many more functionalities in the FPGA. We are implementing both the fronthaul connectivity, and the acceleration capabilities. Of course, the power and the cost of an FPGA is much more than the eASIC.

So, each vRAN vendor needs to decide themselves what they are looking to do. If they are looking for more software solutions, eASIC, the ACC100 is the right solution. If they are looking to create more functionalities within the Layer 1 or accelerate the Layer 1, FPGA can provide a good solution. So, it means that there is no definite answer, right or wrong over here, it's just a decision which direction to go in order to implement the Layer 1 acceleration.

Both are good. Again, performance, price, power, and of course, functionality.

Lilian Veras

That's great. Thank you, Oren. Well, thanks to both of you, Alber and Oren, for the great presentation today. Thank live audience also for joining us today. Please do not forget to give our team a rating for the live recording.

And this concludes our webcast. Thank you all so much.

Oren Benisty

Thank you.

Alber Wu

Thank you.

Lilian Veras

Bye everyone.

Alber Wu

Thank you. Bye-bye everyone.