GaN and its Unique Potential in Communication Systems

Nitin Dahad (00:05):

This is A Smarter World podcast, focusing on the technology and issues behind today's connected world. I'm host, Nitin Dahad, editor at EE times and Embedded.com. In this episode, we'll discuss the semiconductor technology known as gallium nitride with Paul Hart, Executive Vice President and General Manager of radio power for NXP Semiconductors. Gallium nitride, often known by the periodic table name of GaN, is a material that is increasingly in demand in high power and high frequency devices. We'll discuss this technology and a new NXP fab that opened recently in Chandler, Arizona with Paul Hart. Welcome, Paul.

Paul Hart (00:41):

Thanks. Good to talk to you.

Nitin Dahad (00:43):

EE Times covered the opening of a new NXP fab that opened in Chandler, Arizona. We'll go into that in a minute, but I'd like to start at the beginning, as I say, with the basics of gallium nitride and its potential uses. Can you give us a bit of backstory? What is gallium nitride, at a basic level, and how did it find its way into the semiconductor arena? Then, we'll talk a little bit about what kinds of applications, but let's start with the beginning. Tell us a little bit about gallium nitride.

Paul Hart (01:07):

Well, gallium nitride is a wide band gap semiconductor technology. It's comprised of gallium and nitrogen, as the name suggests. And it's a technology that has about three times the band gap of what you'd find in common silicon. So, this very wide band gap technology allows the device to operate at much higher voltages and higher temperatures than silicon without breaking down. And so, it enables higher performance without compromising reliability. GaN, as a whole, also has a much lower on resistance enables the technology to deliver higher current than silicon could.

(01:44):

So, you have a technology that supports both higher voltages and higher peak currents, and therefore, it can deliver much higher power in an RF application up to an order of magnitude higher power than silicon. It first found its uses in the semiconductor arena in very high power radar applications. The technology really began its journey, I would say, in the early '90s in a laboratory environment as an experiment looking at ways to generate higher power densities and higher band gaps. And over time, it grew and found its home first with radar applications at high voltage jamming systems, and now, is having strong inroads into the communication sector.

Nitin Dahad (02:25):

I just said it was niche and very [inaudible 00:02:26] applications. What kind of applications would it have evolved into? I know you talked a little about communications. Maybe just identify maybe some of the areas in which you can use GaN, for example, over silicon.

Paul Hart (02:38):

GaN takes a variety of uses. What we're focused on is in the RF space. So, we're focused on GaN power amplifiers for communication systems. Those same power amplifiers, obviously, can be used for radar applications at high frequency. They can be used for broadband noise generation, as I mentioned, which are very much aerospace and defense applications. So, pretty niche applications. Communications is, obviously, a much bigger market, and it's a bigger focus for NXP. Beyond that though, gallium nitride, particularly, GaN on silicon has strong inroads in the power space as well. So, power conversion, where GaN delivers higher switching speeds than silicon, and therefore, can enable higher efficiency for DC-DC power converters as well as higher power, given the high breakdown voltage. High power DC-DC systems in the 600-volt range are very much a sweet spot for GaN, as well as very small power supplies. So, you see the bricks that come with consumer electronics get smaller and smaller. A lot of that is enabled by GaN on silicon, which has very high switching speeds, and therefore, requires less space and fewer capacitors and inductors to make the system work.

Nitin Dahad (03:49):

Would it be fair to say, NXP's not a newcomer? How long have you been involved in gallium nitride?

Paul Hart (03:54):

We're definitely not a newcomer in this space, but we've had a very specific focus on the communication segment, on the RF space in our investments. We'd started our journey over 20 years ago with initial work first through universities, and then, through internal laboratories. And over the last 20 years, we've really matured the technology, matured the process capability, the IP, et cetera, to create what we believe is a really tailor-built gallium nitride process for communications, something that should work very well in the coming 5G era that we're pushing hard for today.

Nitin Dahad (04:30):

Let's go to the fab, the idea behind the fab in the market. When I interviewed you for the story we were doing for EE Times, you talked about quite a significant investment over the last three years. What was the driver and why now?

Paul Hart (04:44):

We have been a strong player in the communications infrastructure sectors for the last 25 years. We've been a strong player with silicon LDMOS technology, which is the basis for our RF power amplifier products. And that's the technology that we developed in-house, we manufactured internally, and we really engineered or continuously improved the technology over the last 25 years with newer generations, newer nodes being released every 18 to 24 months. So, it's really a continuous improvement process that we've been on, a journey we've been on for quite some time. In parallel, we've been investing in gallium nitride development, realizing that it does have performance benefits over silicon, particularly, higher frequencies, and really wanted to arrive it to the same maturity level as silicon before we released it into our product portfolio. Over the last number of years, as 5G has really started to grow globally, we're seeing much higher demand for higher frequency amplifiers, PAs, to support 5G as well as higher power levels needed at these frequencies.

(05:46):

And silicon LDMOS is a fantastic technology, but it tends to degrade its performance as frequency goes up. So, while it works really well at two gigahertz, its performance is much reduced at 3.5 gigahertz where the majority of the global 5G deployments are occurring today. GaN, on the other hand, with its high electron mobility, has a much higher frequency response, and its performance at 3.5 gigahertz is still very good. It has a much higher efficiency and just general capability than silicon at that frequency. So, with the market move towards 5G, it was clear that it was time to invest in the manufacturing aspects for GaN. And so, we started this investment in 2017, and we've invested since that time, including investments occurring here in 2020 and first part of 2021, about $100 million to upgrade, renovate, re-conceive a site here in Chandler, Arizona to manufacture gallium nitride at high volume, high scale.

(06:46):

And what we want to accomplish is that same tight link between the technology and the end application, and realizing it's, again, going to be a continuous evolution process for our GaN technology over the years to come. We wanted to have that capability in-house and be able to really fine-tune a performance so that we can continue to improve the performance of our end products and help these 5G systems globally become both smaller in size and lower in costs, which really is a key to enabling widespread adoption.

Nitin Dahad (07:16):

That was going to be my next question. What are the strategic considerations? But I think you pretty much addressed that unless there's anything more from a strategic point of view. One of the key applications for GaN is in base stations. Can you tell us a little bit about that? Maybe you can delve into a little bit more detail as to how that works and where NXP's focus is in that?

Paul Hart (07:36):

These base stations, as everyone knows, these are the ugly towers that we see outside. And generally, they provide the downlink in any of these communication systems. In these base stations, you'll generally have a digital processing unit, and you'll have a radio processing unit. In the 4G, 5G era, the radio processing unit sits at the top of the tower, and GaN provides the RF power amplification, so, the final stage of amplification between the RF signal and the antenna. The GaN devices that we're creating, they sit at the top of the tower in these antenna arrays.

(08:10):

In my mind, they're a key enabler to making these boxes smaller and lighter because they're more efficient. And then, allowing, if you will, more boxes to sit on the same tower, which is key, I think, for everyone, because there's one thing we can agree on is that we don't want more towers in our community. But if we can get better utilization out of what's already there, more data throughput with less congestion, then, really, it's a good step forward. So, GaN is an enabler to increasing the data rates of these towers without the need to add more towers to the community.

Nitin Dahad (08:42):

From the top of your head, are you able to give us some scale in terms of how many X times better performance, or better power output, or efficiency?

Paul Hart (08:53):

Yeah. If we look at 5G at 3.5 gigahertz, the systems today are able to deliver five, sometimes, up to 10 times the data rates of 4G systems at lower frequency. A conventional 4G radio operating at two gigahertz can do, to your end device, 100 megabit per second data link, and a 5G system at 3.5 gigahertz can do 500 megabits per second, perhaps up to a gigabit per second in peak applications. 5G also has a different application or extreme bandwidth application at millimeter wave where the data rates can be much higher, north of one, two gigabits per second. And that's, perhaps, an area where GaN will go in the future. But today, its focus is really in these mid-frequency ranges and these coverage and capacity solutions we see being deployed today. But the bandwidth is much higher.

(09:41):

What goes up with that is, of course, the power level that has to be radiated. Because in order to achieve the higher data rates, wider spectrum needs to be used. And if you need to maintain the same capacity and assume the same coverage area, then you have to transmit the same power per megahertz. And so, if you're transmitting over more frequency, that means that the transmit power has to increase. In a 4G radio, you might have four times 40 watts of amplification. So, four antenna streams, each delivering 40 watts in the tower, so that's 160 watts per radio system. In a 5G, you might have 64 antenna elements, each delivering five watts. So, 320 watts per radio system. So, really, a doubling of the output power in order to support the increased frequency and increased data rates. And as such, if we're doubling the output power, we want to not, of course, double the total current consumed. And it makes sense to move to a higher efficiency technology, which is what GaN really is.

Nitin Dahad (10:40):

Got to get even more with 6G, but we'll come to that later. Let's move on to the fab and the background. The fab in Arizona, I think you already alluded to the fact that you're doing some work in Chandler there for a while. But what's going on in Arizona? It's been a hotbed of automotive development. I've read quite a lot about the autonomous testing, for example. And there's been a lot of work in that area. What about the ecosystem for supporting, again, fab and how has that emerged in the Chandler area?

Paul Hart (11:08):

Well, Chandler's a suburb of Phoenix, and semiconductor industry started in Phoenix, I think, in 1949 when Motorola moved to the area, really looking for dry air to manufacturer in. And the Phoenix area being a desert certainly has dry air. The catalyst since then has grown. And what's developed over the last 70 years is a really mature environment for semiconductor manufacturing. Intel has a huge presence in the area. TSMC has just announced that they're building a new five nanometer fab in the area. Other companies such as Microchip and ON Semiconductor have a strong presence here as well. And so, in the Chandler area, you have the ecosystem already established for semiconductor manufacturing. You also have a strong ecosystem for communication systems. Because Motorola, which was a strong leader in communication systems for 50, 60 years, moved to the Phoenix area in the late '40s, early '50s, and really invested heavily.

(12:05):

And so, you have a lot of this radio design, radio talent in the area as well. When we're looking to build a fab where we're focused really on RF applications and communications as a whole, this community really had the pieces we were looking for. It had the talent pool. We had a facility that was already enabled for silicon manufacturing, and it was an ability to just build it out for GaN as well. And then, we had the support ecosystem. So, all the tool vendors, all the ASML, et cetera, those people are all local as well. And they can certainly support the bring up in a pretty fast and comprehensive way.

Nitin Dahad (12:41):

Would you say that Chandler... Because NXP has locations all over the world, and I'm sure there must be expertise elsewhere as well, Chandler had more of that or was there something else?

Paul Hart (12:50):

Yeah. It's a bit historical as well. In the past, we had a gallium arsenide factory. Gallium arsenide is another compound semiconductor, and it's used very heavily also in RF applications, generally at lower power levels. It's the prime technology used in smartphone power amplifiers, so handset applications. We had a large fab here from the early '90s through 2008, and it supported a lot of the Motorola. Then, Freescale products which went into many of the Motorola handsets. When that unwound about 10, 12 years ago, many of the process engineers, engineering experts moved to the silicon factory that we have here in Chandler as well. And they were there all along.

(13:28):

And so, really, Chandler is where we started our GaN technology development. It started, first, when it moved from universities to labs, it moved into that gallium arsenide fab that was here in town at that time. And so, a lot of the knowledge base, a lot of the industry experts were local and still employed. And so, when we looked across the globe at all of our footprints, we found that we had the talent pool, much of it, already in the company, as well as the ecosystem around it, to go and invest here.

Nitin Dahad (13:52):

Right. Okay. Let's now go on to the fab itself. With the 5G GaN opportunity, GaN transistors have a higher frequency and a power density from a fab point of view. What does that get you?

Paul Hart (14:03):

Power density is key. If we have higher power density, we get more interesting metric. The way I think about it is more watts per wafer. And therefore, I need to produce fewer wafers in order to support the number of products that need to be out on the market. So, power density certainly leads to the utilization or the volume that we need to produce. And it's a key attribute in terms of sizing a fab, determining just how much capacity, how many tools we need to buy, et cetera.

Nitin Dahad (14:30):

So, if you're doubling the amount of power you need, you need that better higher density, obviously. So, you're talking about, you're going from 160 watts to 320 watts, for example.

Paul Hart (14:38):

Right.

Nitin Dahad (14:39):

What are the fab's wafer size capabilities? And again, why is this important and timely?

Paul Hart (14:44):

The fab we built supports 150 millimeter wafer production. In terms of a silicon factory, this will be a very old fab, very outdated technology. Most silicon production is either 200 millimeter or 300 millimeter. In the case of GaN, though, we're building GaN on silicon carbide. Silicon carbide is the substrate. It has a very high thermal performance capability. It's incredibly hard substance, almost like diamond, if you will, in its makeup. But as such, it's capable of dissipating and spreading heat very, very actively. And I think what we find is the highest performing gallium nitride in these RF applications. So, we really focused our investments around this high-performing GaN on silicon carbide. And today, six inch or 150 millimeter wafers are very much the state-of-the-art for silicon carbide for RF. And so, we've built our investment around this 150 millimeter diameter. Most GaN for RF is actually built on four-inch or 100 millimeter diameter wafers. And so, we've already taken a step towards the very high end capability, already, for gallium nitride. And then, we built our investment around that, around that wafer diameter.

(15:54):

In terms of tooling, we've invested for very high frequency, high performance GaN. So, we put photo tools capable of driving sub 10th of a micron gate length for GaN, which will allow GaN to scale up in frequency well into the millimeter wave range, essentially, in the 28 gigahertz range where we see the enhanced mobile broadband applications of 5G today. And then, certainly, into the 6G era and beyond. So, we've really geared for high volume GaN on silicon carbide manufacturing, frequency scalability towards 6G in the next 20 years of applications. And then, we've also layered on the automotive heritage of NXP. And so, we've taken all the best auto quality systems, manufacturing processes that we use in our other internal factories, which have to support 100% reliability application, zero parts per billion failure, et cetera. All those tools, because they already existed in the company, have been applied in this gallium nitride factory as well. And so, we're very confident that will result in extremely high quality GaN and something that will become, really, a benchmark technology for the industry as a whole for the years to come.

Nitin Dahad (17:03):

Is there a particular NXP GaN flavor, or is there something different or significant what you do, and how you develop GaN technology NXP compared to, maybe, somebody else?

Paul Hart (17:13):

Sure. I mentioned previously that GaN really began its life in these aerospace and defense applications. Radar systems are a prime example. In a radar system, what you're looking to do is, obviously, amplify a signal, and then, you're sending out chirps signals. Transmit and receive, wait, and use that as a means of detecting things remotely. In GaN, as it was initially engineered, was excellent at that. And that's where it found its first home. When you try to take that same technology and you transmit much more complicated signals, high order, digitally modulated signals, which are used for communication systems, what you find in a lot of GaN is that you have an issue known as memory effects, which basically means that there's a delay in the response of the output signal to the input signal. And that comes from the technology itself.

(18:01):

And in many of these aerospace and defense applications, that was a non-issue. But in communication systems, it creates quite a few difficulties. Right? Yeah. You can imagine that really undesired when you have to transmit a lot of data very fast. So, what we've done is we applied our background in the communication sector, and we really focused, in the half of the last 20 years, all of our investments towards addressing these memory effects issues and developing a very high linearity, low memory gallium nitride process. We think that's unique in the industry. We think we've done it better than others. Obviously, our customers will be the ultimate judge of that, but that's really where we chose to focus our innovations. And we look at all of our GaN or we characterize all of our GaN in these very high complexity, 5G applications with all the digital modulation, all the digital correction that's required in these communication systems already in place, is we optimize our technology. Then, that's something we think we do uniquely in the industry.

Nitin Dahad (19:02):

Finally, on the technical side of the fab, I think we already talked about where you're placing the GaN. It's on silicon carbide. Did we talk about the process node?

Paul Hart (19:09):

No, we didn't. We're developing a pHEMT technology, the initial devices out of the factory now, they have a gate length of 0.4 microns, which when you're talking about five nanometer technologies for silicon, it's quite a ways off. But what that does is it allows us to generate a lot of power at the frequencies of interest. The subsequent variants of GaN are our roadmap, very much drives the gate length down. And as we do that, the peak power capability of the device will reduce, but the frequency response will increase. And so, layering in will be steps along the way of quarter micron GaN, 0.15 micron GaN, sub 100 nanometer GaN. And what that'll do is it'll, of course allow us to continue to address higher and higher frequencies. So, millimeter wave spectrum, or 6G. And you will see the peak power reduce as a result, as we move up in frequency. So, the voltage that we operate at will come down a little bit as frequency response goes up. But it's very much a focused journey and really, steps that we'll take as those areas of the comms market open up.

Nitin Dahad (20:14):

That leads us very nicely onto 6G. I visited the University of [inaudible 00:20:18], and they were doing the 6G research. And when I saw the number of antennas on an array, and it was crazy, and the amount of power you need. Let's close on 6G terahertz communications. That's a way off. And I guess, what do you see in terms of where GaN can take you towards that sort of high frequency communication? We're currently sub six gigahertz for 5G, but I'm guessing it will go more. Tell us a little bit about your thoughts on all of this.

Paul Hart (20:44):

Well, my thoughts are that if we look at 5G today, it covers everything. 5G includes coverage solutions at 600 megahertz with very high radius cell sites, all the way to 28 gigahertz, 39, 47 gigahertz, enhanced mobile broadband, or fixed wireless access applications with just extreme data rates. One, two, maybe up to five gigabit per second data rates in some of these peak applications. GaN fits very nicely into that mix, really in this mid band spectrum today where we're operating at frequencies that are higher than were used in 3G and 4G applications. So, really, north of two gigahertz and that 2.5 to 5 gigahertz range. That's really a sweet spot for this GaN variant that is out in the market now. And over time, while the millimeter wave applications are typically addressed by other technologies today, technologies like silicon germanium, have a much stronger presence at those frequencies than they do now because they lead to monolithic integration and much easier to create a dense antenna array with a technology where you can generate four, to eight, or 16 antenna elements on a single chip, where GaN would be much more of a discreet assembly.

(21:56):

But over time, when we look from 5G to 6G, what we'll see is that all the existing spectrum that is out there today will continue to get used. It's still much of the best spectrum, particularly for coverage, but the higher data rate applications will move up in frequency as well in order to access more bandwidth. So, the mid band technologies we see here at 3.5 gigahertz may move to 10 or 12 gigahertz in the 6G era. The millimeter wave applications, the 28 gigahertz, will likely move to 100 gigahertz or maybe even 140 gigahertz applications where much higher swaths of bandwidth are spectrum are available, and therefore, much higher data rates can be generated. Over that period, I see GaN moving up in frequency, obviously. I see it continuing to be the workhorse technology in this mid band spectrum today and the new mid band spectrums tomorrow, up to 15 gigahertz.

(22:46):

I see millimeter wave GaN coming into play, particularly, in this 28 gigahertz domain. And I see the most intense data applications really being delivered at higher frequencies, likely with still a silicon germanium integration. But over time, GaN can play a role there as well. And so, it's very much a migration, I would say, over time, just frequencies continuing to go up and power levels needing to increase to support the higher data rates. And that's where GaN will fit.

Nitin Dahad (23:13):

Thank you very much. It was a very interesting discussion.

Paul Hart (23:16):

All right, well, thank you, Nitin. It was a pleasure to talk to you today.

Nitin Dahad (23:19):

This has been The Smarter World podcast with me, Nitin Dahad. Thanks for listening, and see you next time.