Podcast Transcript: Foundry Capacity Expansion Easing Chip Shortage, But Weakening Demand Can Cause Oversupply

[00:00:10] **Peter:** Hello everyone. And welcome to The Counterpoint Podcast. I'm your host Peter Richardson. Today I'm joined by two of my analyst colleagues to discuss semiconductor manufacturing. Now, first up we have Dale Gai. Dale is one of our Research Directors who oversee the semiconductor and Foundry research.

Dale is based in Taiwan. Hi Dale. How are you today?

[00:00:33] **Dale:** Yeah, I'm doing good. Thank you, Peter. How are you?

[00:00:35] **Peter:** Yeah, I'm very good. And we're also joined by Senior Analyst, Ashwath Rao. Ashwath is based in India. Now Ashwath holds a doctoral degree in microelectronics from the Indian Institute of Information Technology and Ashwath covers Semiconductors and Component Research for us here at Counterpoint. Hello, Ashwath. How are you today?

[00:00:55] **Ashwath:** Yeah. Thanks, Peter for having me on the podcast. I'm good, Peter.

[00:00:58] **Peter:** Great. So now since the start of the COVID-19 pandemic, semiconductor components, and particularly shortages have been increasingly in the news, you know, in addition to the COVID-driven disruptions, there is also trade tensions between the US and China that upset the normal supply situation.

Now, of course, we have the Russia-Ukraine conflict, which has added further uncertainties around, particularly some raw materials that are used in semiconductor manufacturing. Now, these issues have caused various shortages in multiple industries. Perhaps the automotive sector has been the highest profile casualty, but now we had this, the supply situation seems to be improving.

And now we're sort of getting into a situation where Macroeconomic factors are impacting on demand. So we're potentially, sort of, having the situation flip. So let's, let's get into the discussion a bit and we'll try and unpick some of these issues, what they mean for component manufacturing today, and then kind of looking out a few years.

So, Dale, let me start with you first. Now, as I mentioned, there's been this. Supply-demand, imbalance leading to shortages of some components, but now demand is dropping off. You know, we see sort of lots of macroeconomic factors, inflation rising, for example, you know, input costs for many manufacturers are increasing.

So can you explain a bit more about the foundry inventory correction cycle that we're kind of going through and how that's impacting on various products and how you see that developing over the next couple of quarters.

[00:02:31] **Dale:** Sure. Thank you, Peter. Yes. The inventory level has been increased significantly since early this year. And we start to see a sharp declines of demand of chips in both the wafers and also the end products in consumer electronic segment because of the high inventory and all the macroeconomic headwinds you mentioned. For example, we access the data smartphone AP inventory, which is currently made of the application processors for smartphones, the inventory days to be 65 or even to 70 days by end of 2022, if we compare to 45 to 50 days as a normal pattern in the past few years. And the high inventory also happens in many other smartphone components, including PA chips, the power amplifier chips in the RF modules, some driver ICS for panels and CMOS sensors also in the smartphones. So relatively we think that will be pretty end of bad of inventory corrections in the next few quarters. On the other hand, relatively we see healthier inventory levels in automotive and industrial goods, such as power management, microprocessor chips, and FPGAs. That because the industry still in high demand of dollar content increase of the silicon chips and also the automotive supply chain, still building up the inventories rather than correction of the inventory cycle.

If we look at historically the inventory adjustment cycle, usually it takes about five to six quarters to back to a normal level. So if we have the inventory right now we probably see better on the balance of more healthier inventory level improve likely in the mid of 2023. That's our current view on answer these question.

[00:04:27] **Peter:** Okay, but what we've also seen over the last year, or two is big spending intentions in terms of Capex by foundries. So they responded to the, the supply disruption situation, sort of in 2020-2021 making promises to spend huge sums, billions of dollars on kind of creating new Capex.

And some of that was in, I guess, in the leading edge nodes, but also in some of the more mature nodes. Right. So can we expect that this will lead to a big increase in capacity in a year or two that might lead to an oversupply situation? How should we think about that?

[00:05:04] **Dale:** Right. Thanks. If we talk about the capacity utilization rate, which is a good indicator to see the demand and supply balance, I would say is case by case. If you look at TSMC, which they just report a very strong quarter in Q2, quarter two and also expect over 30% of annual growth rate in this year. Even with inventory concerns, we still expect fully loaded of capacity in advanced node, such as 7nm, 6nm and 4nm to 5nm in a second half, driven by Apple and driven by lots of the HPC high performed computing products.

The mature product line could be some slight decline, maybe utilization rate back to 95%. But overall, these situation is still very strong. However, for second-tier foundries, if they compete pretty much-commoditized products and highly tied to consumer products in both 12-inch and 8-inch fabs, the utilization rate might drop to 90% or even below.

So overall we think over-supply issues are likely to happen in a more mature node due to the lack of entry barrier in capacity expansions. Just as you mentioned the China and also we see some Taiwanese build of the mature node capacity in the past one to two years. And that should be, be the issue if that then disappoints.

Particularly if we're using 40 to 45-nanometer example, they're high running a higher risk of oversupply in the next year, because a lot of applications such as driver IC sensors are more applied to smartphones, you know, and TVs where we might not expect meaningful recovery of demand. And the 20 nanometer, for example, this is a lot of inventory capacity expansions this year.

The MLA driver IC might have some corrections, but the demand from Wi-Fi, the demand from automotive MCUs there will be offset. So it's mixed specs. But overall we conclude the mature nodes might have more oversupply risks than the advanced notes, which be pretty dominant by the very few top players in need of the high performance computing chips.

[00:07:20] **Peter:** Great. Okay. That's really clear. Thanks Dale. So let me let me bring Ashwath into the conversation. Now, as Dale mentioned, foundries are spending large on new capacity which means new manufacturing equipment. So can you explain the role of the wafer fab equipment manufacturers in the overall supply chain and what are the driving factors we should think about in relation to that?

[00:07:43] **Ashwath:** Sure Peter. So I'll just start with understanding the semiconductor production equipment and then move to the major players and followed by the dry factors. So semiconductor production process is divided into front-end production in which circuits are found on the wafer and inspected and backend where the wafers are actually cut into chips and inspected again.

So ideally the equipment ranges from deposition, etched systems, lithography cleaning systems, you have metrology and inspection also followed by advanced packaging and advanced process control. So top five WFE players contribute to roughly around 80% of the market. You have Applied Materials, a leading manufacturer in etch deposition and chemical mechanical planarization, ASML leader in advanced lithography systems with their deep ultraviolet and EUV systems.

Tokyo Electron is also having a significant market share in etch and deposition systems. LAM research is an innovative fab equipment manufacturer with competency and deposition, etch and clean again. And KLA is a global leader in process control equipment, providing process enabling solutions for manufacturing wafers.

So WFE market in 2021 has been outstanding with an increase of around 33% year on year, and primarily driven by strength across the segments, NAND DRAM, and Foundry logic. And major driving factors are critical technology transformations, increasing device and manufacturing complexity, aggressive investments for higher semiconductor performance, and these will be driving the WFE spending in 2022.

[00:09:23] **Peter:** Got it. Okay. Thanks. Ashwath. Now there are various geopolitical factors in play. I mean, we have the Russia Ukraine conflict, of course. US-China trade tensions, where we're seeing some potential kind of decoupling of manufacturing from the traditional kind of geo basis where they've been built up over the last decade or so. And you're seeing this, I think in the forthcoming chips act, which is coming through the US government right now, there's some question marks about whether it'll actually be put into law before the summer recess and then the midterm elections. But if it goes through that could see a resurgence of US-based chip manufacturing, which is seen as strategically important by the US government. So how should we think about these issues in relation to lead times for some of this fab equipment, which is, expensive and highly complex as you've outlined, not very many players able to supply this kit.

[00:10:19] **Ashwath:** Sure Peter. So semiconductor shortage was initially triggered by the pandemic. Then geopolitical tensions broke out. Then again, one more round of COVID and lockdown and just entirely choked the supply chain. So now it has become more complicated due to the lack of tools required to make the chips. And that's the equipment makers need chips for their equipment also.

So for example, the Russia-Ukraine. Right. The critical materials needed are palladium and neon gas, where palladium is critical input in the metal connections, attaching chips to the circuit, and neon gas is critical for the examiner lasers that is used in deep EUV lithography operation. So, and most of the materials are source so strong Ukraine and Russia.

However, the requirement of neon gas in enabling DUV lithography equipment also has been significantly reduced due to the improvements in technology. And yes, long-term conflict has increased the lead time and delayed delivery of Wafer fab equipment, which will be affecting the foundry schedule also. So the semiconductor equipment supply shortage is spreading in all the direction from the front end process to the back end process.

And one more thing to notice the wave of manufacturing equipment are unique to each company and it's very difficult to have alternate sourcing also, and further, the verification process is very lengthy to ensure production continuity and quality also is of utmost important. So yes even the OSAT industries will have to face delayed shipment of equipment which are exceeding more than a year.

The lead times are greater than a year and the equipment manufacturers are currently prioritizing to address the demand-supply gap of front-end process equipments. And later on, they'll be moving to the backend equipments. Secondly, governments also around the world have planning to spend billions of dollars to protect against supply chain, disruptions, and globally it was an integrated rated supply chain, but they are each government wants to localize themselves. And one more thing, the US government is implying restrictions on the ban of EUV and DUV systems, deep lithography systems to China. However it could aggravate the component shortages in the future.

[00:12:32] **Peter:** Yeah, it's interesting. So we're seeing a, kind of a parallel tracks almost emerging between, you know, China and the US on the manufacturing side. So if I can bring you back in Dale, can you tell us a bit more about how you see the capacity expansion developing relative to these potential equipment delays and how the capacity is likely to develop over the coming few years? I think you outlined earlier the initial phase over the next couple of quarters, but looking longer term, what's your expectation about how the industry develops from a longer-term cycle perspective?

[00:13:06] **Dale:** Right. Thank you, Peter. I would like to have from two perspectives. Number one is about the equipment delays, because I agree that the pace of the new capacity some of the foundries and IDM slow down in second half and maybe 2023 in light of the demand weakness and however, on the flip side different from the previous cycles, the supply of the equipment tools also constraint as a good excuse as a Capex adjustment. TSMC adjust down its Capex in 2022, slightly in July conference. Not because they're turning conservative, but they saw the delay of the delivery in some equipment tools. So while we still concerned about a demand slowdown, a risk of oversupply, but we probably see a better supply an angle because of the some equipment to delay.

So to alleviate the further oversupply risk into the next year. And however, if you look at the longer-term perspective we think is subject still subject to drivers on a demand side, it's up to the high-performance computing and data center, edge computing, and autonomous driving. And these will supply, sorry these of demand will support the growth of the advanced technologies. On a supply side definitely we need to see an execution of the new green facts especially on the fabs subsidize my government, such as the US. So we still highlight that there's still a lot of the risks in the future on if the demand of the supply come to the, in a line which faster than expected.

However, in this industry, we have to highlight that there's still lack of engineering talents running the fabs globally and keep in mind, the supply of key EUV tools and materials are quite limited. So we still too early to tell the long-term perspective while we know next year to be a bit of oversupply cycle.

But we still need to see that if the demand for the advance now continue going up. We actually, we still positive on the semiconductor correction cycle after this inventory duration in the next few quarters.

[00:15:23] **Peter:** Hmm. Okay. So yeah, really interesting. And, and actually that's is a kind of a good lead into last part of the discussion, which is to look much further ahead. So currently, I think the leading edge process node is at 4 nanometers and TSMC is. Testing 3 nanometer already. And it's likely to start production before the end of this year. And I think they've then spoken about 2 nanometer coming in 2025. So coming back to you Ashwath.

So thinking about how this develops over the next few years. I mean, is, is there a limit to how far we can shrink the, the line widths? I mean, are we gonna get to a point where we're measuring line widths in angstroms rather than nanometers? How far does this go?

[00:16:07] **Ashwath:** Sure Peter. Yeah. So as device scaling continues to go below 3 nanometers. So processing of each of these modules in the front end, middle end, and the backend becomes very challenging. So just to understand the FEOL MOL and the BEOL, so Front End of Line (FEOL) are basically the transistors which are being scaled down. So currently FinFETs are at the FEOL, and they're reaching their limit. And the Middle End of Line (MOL) is basically the contacts and due to the increase in the source drain at contact resistance, that's also another issue that has to be addressed.

And in the Back End of Line (BEOL), basically refers to the interconnects on the top of the tip and routing congestion and RC delays also have become important bottlenecks, which has to be addressed. So if you talk about Moore’s Law four critical metrics, basically Power, Performance, Area, and Cost, and power and performance have already fallen of the metrics from Moore’s Law.

However density. Yes. It's able to follow Moore’s Law, for example, TSMC 3 nanometer there's an improvement in density 1.7X times improvement over the fine process, but yes, cost component it's increasing drastically. And in order to enable the close to Moore’s Law chip scaling, lot of new transition architectures including new materials, improving the patterning of interconnects patterning of the contact resistance design technology co-op optimization, advanced packaging. All this has to be adopted by the industry. So once the FinFET reaches the wall, yes, chipmakers will how to go to nano FETS and Gates All Around.

And Samsung has already started producing 3 nanometer with GAAFET architecture. Recently IMEC the most advanced semiconductor research firm in the have shared their sub-1 nanometer silicon and transition roadmap at their summit. And they'll mentioned that yes, extension of Moore’s Law is visible till 2036. So the innovations at the front end could be GAAFETs followed by Forksheet followed by complimentary FETS in MOL it could be scaling boosters, such as buried power rails or self aligned gate contacts. So buried power rails are nothing, but you deliver the power and signal lines from the bottom of the chip, and then different materials are being used in context, basically. And in back end of line it's hybrid metallization and semi damascene process.

So, yes, Intel 20 Armstrong with GAAFET EUV in high volume is expected to be coming out in 2025, followed by 18 Armstrong and then it could be either 16 or 14 Armstrong in 2028. And TSMC N2 is expected to be around in high volume manufacturing by around 2025 but with GAAFETs and 1.4 nanometers around 2030.

[00:19:02] **Peter:** Great. That's really interesting Ashwath. Thanks. It sounds like there's still some road ahead for the progression of, of Moore's law in some shape or form. But a lot of what you just talked about sounds like there's gonna be quite big implication for the semiconductor manufacturing equipment vendors. So are they also gearing up to support these developments, particularly in terms of different packaging technologies?

[00:19:27] **Ashwath:** Yes, Peter. Yeah. So historically the primary function of semiconductor packaging was to protect the dye and connect it to PCBs. But now in the era of flight computing, both the chips and the systems have taken center stage. So heterogeneous design and advanced are completely imperative for the leading semiconductor companies and the recent announcement also to form a consortium for UCI that's a universal chip interconnect express is a step forward to stretch further beyond transited scaling by utilizing advanced packaging techniques. Again, the next revolution could be hybrid bonding also where the chips and the wafers can be directly connected in copper with no need of microbumps resulting obviously increased I/O density and improvement in energy androgen normally allow the engineers to disintegrate a monolithic design into smaller chips that can be connected in a single dye. So they have a engineers can mix and match the performance critical dye with other chips. We also in the recent earning calls, Q2 earnings call said we'll be implementing the plate platform, scaling concept, wherein benefits of power delivery schemes, advanced packaging and chiplet will be utilized to control cost and have an overall advantage. The entry node also will be the longest node to be used before migrating to N2. And they'll come up with the TSMC FinFLEX architectural innovations, which offers flexibility to the customers to create designs precisely tuned for their needs with functional blocks implement.

So going forward in the future, the packaging different packaging technologies will be fan in, in wave level packaging for mobiles, it could be Flip Chip package on package. For CPU and GPS it could be 2.5 3D extended to Flip Chip BGA. For RF it could be antenna in packaging, and automotive again, it could be Flip Chip ball-grid array packaging.

[00:21:25] **Peter:** All right. Sounds really fascinating. So just before we wrap up Dale, any last thoughts from you on how you see the kind of the longer term developments in the semiconductor cycle? You know, looking out, I mean, Ashwath was talking about 2036. Any thoughts from you on the, that kind of long-range per perspective?

[00:21:44] **Dale:** Yeah, actually I'm quite excited about the technology continuing to migraine to the next generation of the chips architecture. And we see a lot of demand. I believe lots of the demands such as the hyperscalers, lots of the new chips the AI applications, autonomous driving, and even these wearable AR/VR we haven't been there. So I think there's still a lot of applications where the leading foundry such as TSMC is developing these products with their key clients in most advanced notes, applying the most of the technologies and equipments materials.

So if you look at a cycle for this year, definitely a little disappointed, especially the consumer and inventories, the headwinds of global macro, but I believe this a technology continues to migrate and we think the foundry and WFE industries still have a lot of the upside to go in terms of the new product development. That's all my comments today.

[00:22:46] **Peter:** Thanks, Dale. Yeah. Fascinating stuff. And thanks. Thanks also to you Ashwath really interesting discussion.

[00:22:53] **Ashwath:** Yeah. Thanks Peter.

[00:22:55] **Dale:** Thank you, Peter.

[00:22:56] **Peter:** Great. And thanks everyone for listening to this podcast. If you'd like to know more about our semiconductor research, please do reach out to us. We'd be happy to talk to you. And we look forward to sharing more insights with you on the next counterpoint podcast. So have a great rest of your day.