JOHN VINH: Great. Good morning, everybody. I'm John Vinh, KeyBanc Capital Markets. I cover semis here. We're pleased to have AMD with us this morning and pleased to have Matt Ramsay, Corporate Vice President of Financial Strategy and Investor Relations. Welcome, Matt.

MATT RAMSAY: Thank you, John. And thank you for all your colleagues at KeyBanc hosting us. And I think we got-- I live in Atlanta. So we have a little bit of warm weather here, too. But the humidity is, I think, a factor of six below where I'm used to. So this is great. So thank you guys for having us.

JOHN VINH: Great. Maybe where we could start off our conversation, Matt, is server CPU sounds like it's on fire for you guys. I think you talked about 80% growth in the second half, 70% revenue growth next year. And you've talked about having secured enough capacity to support that growth and potentially even upside to that number. Maybe you can talk through, what's been the primary constraint that you've had to work on to secure that sort of [indiscernible] capacity? And then for the upside to the 70% number, I've got to imagine you've got a lot more than demand than that. What needs to happen in order for you to be able to raise that number going forward?

MATT RAMSAY: No. Now, thank you for the question, John. And it is a remarkable time in the server CPU market. I know there was a couple of years where the server market maybe grew a little bit less than it had historically as CapEx quickly shifted towards AI systems. And we've always at AMD had the belief that server CPUs were going to-- and CPUs in general were of paramount importance across our business and have been investing in multiple generations of CPU architecture over a very long period of time.

We're just about to launch our two nanometer Venice CPUs. We're sampling them to everyone today. They're going to ship in our Helios AI racks. And they're going to ship-- the family of Venice CPUs are going to ship broadly across all of our server markets. What's happened in the last nine months is absolutely phenomenal in the server market. And I think maybe I can describe a little bit about that big picture and then we can get to some of your supply chain-oriented questions.

What we've seen over the last six or nine months is-- I think many of this audience and many in the industry have been waiting to see when the dominance of AI spend around AI training of large models was eventually going to shift towards being much more heavily on inference. And I think at our advancing AI event a few weeks ago put out some models that we've done internally where the shift is happening right now towards inference being the majority of the AI computing spend.

What's happened on top of that shift, and that shift is happening right now, is-- and this is maybe my terminology analogy versus the company's terminology. But at the same time, the shift of spend is going from training to inference. Chatbot inference is becoming agentic inference at the same time. And that's a phenomenal thing for our business because we supply what we believe are very differentiated products for inference on the GPU side given our memory footprint and bandwidth and also the best CPUs in the industry.

So what you need when you're running agentic inference, you'll hand off the tasks to the big XPU or GPU cluster to actually run the intelligence of the inference. And then in a very automated way, the agents will take the result of the prior inference, figure out what to do next and what to ask the inference and the AI model to do next, figure out where do I get the data to support the next step in the inference-- some of it comes from the cloud. Some of it comes from enterprise systems. Some of it comes from the web. Some of it comes from wherever-- reorganize the data and then hand it back off to the next step of the inference.

And you do that a whole slew of times in a very automatic way and you end up with a big agentic automated inference flow. And agents are nothing but what simulated automated workers. And the computing that these workers and agents do is very diverse. I just described some of it, pulling data from here, there, and everywhere to hand off to the next inference task. Sometimes, it's running code that was just generated by the prior inference task. And that requires really high thread count, high-performance CPUs that can do many, many different tasks.

And we're going to push at AMD to make sure that our Helios AI racks are doing as much of this inference computing as we can. But there's a market out there where inference is run on many different accelerators. And all of those need agentic racks of CPUs. And that's what we've seen inflect the market. So we reported that we grew our server business more than 50% in the first quarter of this year. And many of you guys might remember when growing 15% to 20% in the server business was a phenomenal result.

So we grew more than 50% in Q1. We grew more than 70% in Q2. And interestingly, both our cloud and our enterprise business both grew more than 70% in the second quarter. To John's point, we've talked about growing greater than 80% in Q3 and Q4 in the back half of the year. And then a really early view of 2027 is on top of that much larger base, at least 70% growth next year. So the constraints that you were asking about, John, I think there's a few. We obviously need the way for support from our great partner in TSMC.

And Lisa and our supply chain team have been working with the folks at TSMC directly. And they've been absolutely phenomenal partners of giving us additional supply. And we have to actually-- obviously, we need to earn it by delivering the products, but-- then there's advanced packaging. Venice is the first server product in the market to use advanced packaging. And we've invested. Lisa was in Taiwan a month and a half ago or so, and we announced a $10 billion ecosystem investment in the Taiwanese ESOP ecosystem. Much of that is oriented on back-end capacity. So I think we feel really good about where we are there.

And then the industry obviously needs to have memory to support these servers as we sell them. So we're working really closely with all of our OEM and ODM partners and our hyperscale partners to make sure that they have matched set DRAM to support the server volume. So I mean, it's a pretty remarkable stat. The early view of 2027 server revenue for our company is roughly 20% larger than the whole server market was in 2025. So it's a pretty phenomenal thing that's happening, and we're going to continue to innovate.

We talked about not just Venice, but the whole Florence lineup of CPUs that comes in 2028 at our event a couple weeks ago. So we're going to continue to push across really high-performance, single-thread type skews to the highest core count and thread count CPUs for agentic racks and everything in between. But anyway, John, that was a long-winded version, but there's a lot going on in server. So I just wanted to give a little bit of the lay of the land.

JOHN VINH: Great. Thanks. Just a quick follow up here, Matt, is do you think there's opportunities for you to secure additional capacity through the rest of the year to maybe grow it at a faster rate than that?

MATT RAMSAY: Well, we're certainly going to try. As you guys know, there's a lot of things in the supply chain that are tight. We happen to be very well positioned, where depending on what quarter it is, we're the third or fourth largest customer at TSMC and then a very loyal partner with them for a long time. So we're-- one of the things that's been interesting is-- when Lisa speaks about this, the industry is quite good. If you give accurate forecasts and sufficient lead time, the industry is quite good at getting you supply.

It's when you come and ask for stuff way underneath lead time where things get a little more complicated. So we're doing our absolute best to continue to upside on supply and for the second half of the year. But I think we feel-- as we look into 2027 and into '28, the industry, including ourselves and our partners have had much more time to adjust supply higher to support the growth. The nearer term has been where we've been having to work hard to get additional supply.

JOHN VINH: Great. Maybe switching to Helios, you talk about Helios going into mass production with maybe shipments starting in September. I'm just curious, can you talk about just the feedback you've gotten from your customers so far on Helios? What's surprised them the most about it?

MATT RAMSAY: I guess, I would start John that you're absolutely right. We're going to start to ship MI450 and the Venice CPUs and some of our Pensando networking products into our ODM partners that are building Helios starting in the month of September. And then we're going to have a fairly large ramp of that revenue in the fourth quarter, another fairly large jump of revenue in Q1. And the business is going to be pretty phenomenal from a growth perspective. We talked about the server business growing more than 70%, but the data center business, inclusive of AI, growing much more than 100% next year.

So that'll leave quite a lot of-- I don't know what much, much more than 100% is. But the math has to do that to get there for 2027. The feedback from customers has been phenomenal. We've been working with the customers hand in hand on the designs and the spec of Helios for a very long time. And when you get to a product that is as complicated as this one, going and launching full racks with our partners, the goal is no surprises. [CHUCKLES]

And so when you get folks that are running full model code on sampled systems and getting ready to scale those systems and the feedback from them is, wow, this thing really works the way that you told us to it was going to work, that's the feedback that you want. Every day is a little bit of a different challenge, as we're trying to get a product as complicated as this off the ground into significant scale. I mean, we're going to go from a standing start to billions of dollars of revenue in the first quarter that it's shipping for the full quarter. So it's quite a ramp.

And when you have that, there's something every day that happens. But the feedback from the customer base is phenomenal. We're very, very proud to have three of the premier model companies in OpenAI, Meta, and Anthropic as our workhorse customers for this generation of product, each of which want to go to gigawatt scale with our first rack scale solution. And I think that's a testament to the performance of the rack, the performance of AMD as a partner, and also the performance of the software stack that's allowed them to get to that point.

JOHN VINH: Great. Maybe just to follow up on that, you talked about your three strategic partners, OpenAI, Anthropic, Meta. It seems like you've got multi-gigawatt commitments from all of them. It seems like the expectation there is they're each going to roughly deploy about a gigawatt next year. How do we think about that ramp next year? It's kind of a gigawatt per strategic partner the right way to think about it. And then what about the rest of your non-strategic? You've got your core customers such as Oracle and Microsoft. It seems like they could also maybe account for another gigawatt of capacity.

MATT RAMSAY: Yeah, John, I think it's we want to take this ramp. And it's going to be a very fast ramp when you look at the revenue dollars. But we also want to take it in a methodical approach because it is a complicated system. And the focus is on getting stable systems into market that our customers can run production code on as quickly as possible. And so if you think about MI455, the MI450 Helios series being the primary driver of our AI business, that really starts in next month, in September and probably runs through the first quarter or so of 2028, so a six or a seven quarter period of time.

And so we've gotten commitments for six gigawatt arrangements but one gigawatt commitments from both OpenAI and Meta. We've gotten a gigawatt commitment and a two gigawatt ambition from Anthropic over that generation. And as you mentioned, at OCI, there are other OCI customers that are going to be running on Helios. At Microsoft, there's Azure customers and Microsoft's own internal AI workloads that are going to be running on it. There's a number of neoclouds that are there as well. So I think Lisa made a comment a couple of weeks ago that we would love to be-- there's certainly demand there to do what you described.

It's now. It's a matter of us executing and making sure we have land and power and shell and capital commitments for all the folks that actually deploy this stuff. We'd love to be able to do the first full gigawatt with Anthropic in 2027. Whether we get all the way there or not, there's some variables there, but we're on a really exciting growth trajectory to much more than grow our AI business by-- it's going to double in a good bit more than that next year. Whether it's exactly the number of gigawatts per customer, I think that's a little too precise for today. But we have a significant ambition to ramp supply. The customer demand is quite strong across those three buckets and others. And we'll see where we get.

JOHN VINH: That's great. There is quite a bit of excitement from the investor community for you guys in terms of the Helios opportunity going forward. But there's a little bit of angst around this would be the first time you're going to go to rack scale. One of your peers, as you recall, when they first went to rack scale, there were quite a few growing pains. Maybe talk about what are you ensuring? What are you doing to ensure that there's a smooth ramp here into the back half of the year?

MATT RAMSAY: I think the first order answer to that, John, is a lot of work.

[CHUCKLES]

The second answer to that is we've been taking a lot of feedback from the customer base over the last couple of years, as we've been designing and getting ready to ramp Helios and setting up the ODM partners and the supply chain. Third, we did a significant acquisition of ZT systems to bring in system-level talent into the company. And those folks have proven invaluable to hardening the design and de-risking the different design points. There's a very large Gantt chart to ramp a complicated product such as this, as you might imagine.

And every day is a little bit of a different battle. But we've tried to be methodical about it about it. We've tried to design the system such that we've taken risk out of the design. We've been-- we're going to be fairly focused on the initial ODM partners to ramp. We're not going to ramp everyone to massive scale all at the same time. We're going to have a couple of focused partners to start and then spread it out into the ODM ecosystem much wider as we go forward and once we've gotten sufficient scale to copy exact the success we have with the first couple of partners into a broader system.

So I think there's a number of things where, as I said, the first focus is to make sure not that we're just shipping racks, but we're shipping racks that are running code, production code for customers as quickly as possible. And so that-- I guess, the way that I would describe it right now is there's no smoking guns. We've gotten over that. We'll ship sampled racks. People are running code. They're very, very happy with it. We've proven that we can build racks. Now, the question is the vast amount of blocking and tackling that we need to do to build the racks at the scale that we're talking about. That's the next step and where the team is focused.

JOHN VINH: Great. Any questions there? OK. At your advancing AI event, I thought one of the most interesting announcements you made or comments that you guys had made is that on ROCm.ai, that this is probably the biggest leap that you've made from a software perspective. Can you just unpack that a little bit and just talk about what you guys are most excited about there? It seems like you feel pretty confident. You've been able to close the gap with CUDA there.

MATT RAMSAY: No, I think it's a great observation, the team there with Vamsi and his leadership and Anush in the software team on ROCm have done. The progress that they've made in the last 18 months has been phenomenal. And it's been accelerated significantly in the last six to nine months of using AI tools in software development. One of the things that we announced in our multi-gigawatt partnership with Anthropic that a lot of people focus on are maybe the hardware pieces, but some of the software bits are just as important I think.

We're not only using Claude across AMD's engineering teams broadly, but we're working with Anthropic to make sure that any other customers that use Claude for their AI model work can automate and land on top of ROCm, on top of our Instinct platforms their code that's automated by Claude. And I think that's an important step. The Anthropic people were kind enough to tell the story on stage with Lisa so we can repeat it.

But one of the things that they did going back in the really early part of this year is they actually rented a cluster of MI355s and got their premiere inference model up and running on 345 and tuned in a weekend. And I think that gives you some-- we can give you all the kind of stats about ROCm closing the gap with CUDA that you want. But the fact that a premier model company out of the gate without AMD's help, or even our knowledge at that time, can get up and run in a weekend, that was a pretty phenomenal result and shows you where the software is now.

Software is a battle every day. ROCm AI is going to launch with Helios. There'll be another ROCm version that launches with the 500 series next year. And we're always in a refresh battle there. But I think we-- for the largest customers that are spending the majority of the CapEx and are the most sophisticated in terms of their model work, we feel like we've taken the friction out of the system for ROCm to be a great place for them to do their work and as they ramp Helios.

And that gap has, I guess, narrowed to a point where it's not really a conversation with the top customers now. They know what they want to do with their application at that level, and they know that they need to run it through ROCm to get to our hardware. And that's what we're optimizing for. And you can tell with the first generation product that they're using from AMD on Instinct and our first rack scale product running software on top of ROCm that a company like Anthropic has ambitions to do up to two gigawatts with us in the first generation is a pretty good testament to where the software stack is.

JOHN VINH: Great. I thought one of the most interesting things you also talked about is you updated the server CPU TAM to 220 billion by 2030. I really liked how you broke out the key workloads within that stack, with Agentic representing roughly about 2/3 of that TAM by 2030, which is pretty interesting. And I think you reiterated expectations of getting to 50% market share.

When I think about the two camps of competition for you, it's your other x86 peer, and then you got ARM. If you look at the ARM results, it does look like they are gaining share albeit off of a smaller base. Can you just talk about those two camps of competition and how you think about AMD faring against those two camps.

MATT RAMSAY: Sure. I think the first thing I would say, John, just philosophically in AMD, the discussion about what we want to do in the server market does not start with x86 versus ARM. It starts with go build the best server parts. And that-- I think that's the most important piece of this conversation, is-- and do it with huge platform support. And because of the chiplet architecture that we bring with a relatively small number of actually taped-out chiplets, we can make a large, large number of optimization points and SKUs across the server business relatively easily.

So what we're seeing, you mentioned the way that we broke out the TAM. We're seeing the same medium core count, really high frequency up to 5 gigahertz with high bandwidth emerge as a market for AI head nodes, head nodes for GPUs or XPUs, and that market be very distinct in its characteristics relative to the CPU-only racks that run agents that we talked about earlier in the conversation, where that demand is predominantly for our largest core count, largest thread count products that--

I mean, then this goes up to 256 cores and 512 threads. And it's how many agents can you run in a megawatt, or how many agents can you run in a rack footprint? And then we have the enterprise market and the cloud market in between of traditional server workloads. Like these agents will make a lot of calls to CRM systems or ERP systems or databases or whatnot. And those will run either in the cloud or on prem, depending on-- that's really not a workload decision. That's a deployment decision.

And the optimization points for those three buckets are very different. And we're seeing demand pull from customers for different SKUs to support those different buckets right. If you think about one of the advantages that we've had as we've gained x86 share, I can remember when AMD's share was 0.4. And now, it's in the high 40 of the x86 market. One of the advantages that we've had is because we've had so much expertise on this chiplet technology, we've been able to push the SKUs and the core counts very rapidly versus our competitor. And I think we can continue to do that.

If we look at versus some of the ARM competition, again, build the best CPU regardless of instruction set. And some of the x86 security and reliability, availability, serviceability features that we've hardened in our EPYC roadmap from servicing all the enterprises and all the hyperscalers over the last five or six generations, those-- without being put through those paces, I think that's going to be difficult to replicate. And if you think about the importance of having security around running agent codes, I mean, an agent is what? An autonomous worker with access to your enterprise data.

So the differentiation in RAS and security features and the ability to run all of the x86 enterprise workloads in addition to run the agent code, we feel strongly that not only are we going to be participating in a TAM that's much, much larger as you described, but the ambition to get to over 50% revenue share of that much larger TAM is certainly still there. And the indications that we're getting from customers, as we partner with them on what the roadmap looks like for not just Venice, but the Florence generation, the Ravenna generation, are the engagements. There are really, really deep.

JOHN VINH: Last question for me is there's a little bit of debate about which architecture, ARM or x86 is better optimized for agentic workloads? I think you guys said that Venice was CPU that was built for agentic. Obviously, you probably feel that your CPU is superior. What's the key metric that we should be paying attention to that you think suggests that maybe your server CPU is better optimized than ARM for authentic workloads?

MATT RAMSAY: I think the first-- I would make two points to start, John. One is agentic is not a workload. It's a very diverse set of workloads. There's not one design point or one skew point or one optimization point that is going to be the right CPU for AI. I talked a little bit before about the big divergence that we're seeing in characteristics of head node CPUs and traditional workload CPUs and what's going to happen in agents.

And one of the metrics that keeps coming back to us on the agentic rack piece is agents per megawatt or threads per megawatt. So where we're seeing the demand pull on the agentic side is for our highest SKUs and our highest core count SKUs. And so I don't know that it's an instruction set conversation. It's a capabilities conversation. We've been-- and as we ramp Venice, which is the first server part in the industry to use really advanced packaging and have really high core counts, that's where we're getting the demand pull in the agentic side.

So I think the first point is probably the most important, which is agentic AI, from a CPU perspective, is not a monolithic workload. It's a very diverse set of workloads for which there are very many optimization points. And the way that we bring together our roadmap with the configurability that we have through chiplets to address all of those things with a large number of SKUs at scale, I think really does differentiate the roadmap.

JOHN VINH: Great. It looks like we're out of time. Thank you, Matt.

MATT RAMSAY: Yep. Thank you, John. Thanks, everyone.