

Please welcome to the stage Arm's Executive Vice President of Cloud AI, Mohamed Awad.

Mohamed Awad - Executive Vice President, Cloud AI Business Unit, Arm Holdings plc

Feels like déjà vu a little bit. I'm hoping the slides work well for me this time, so let's go through it and see where we're at.

Obviously our opportunity in cloud is huge. We're seeing a tremendous amount of traction in lots of different ways, and I'm going to spend some time walking you through how I think about the business, how I think about the market opportunity, and also how I think about our go-to-market strategy.

How we engage the customer through a couple of different channels now. Obviously one is IP, another is CSS, and then we've got Arm AGI CPU as well. So I'm going to talk you through that.

Let's level set on the size of the market. Like I said, the numbers are big, right? Accelerators obviously make up a huge, huge chunk of the potential opportunity. I'm not gonna talk too much about accelerators today, but I am gonna highlight, and I think it's important to understand that the accelerator opportunity we're seeing in the accelerator space is actually driving a tremendous amount of growth in the other areas as well.

And I don't think that can be understated. You've seen that with the explosion of data center CPU demand recently, and moving forward we're actually seeing that in other markets as well. As these things become more and more agentic, the demand around CPUs continues to grow.

We think the data center CPU market is well over \$100 billion by fiscal year ending 2031. Jason is going to go into some of this in a little more detail later.

I want to show you a different view. When I think about the product view, the go-to-market view, I think about it slightly differently. I think about it first and foremost like the cloud more broadly. The folks that are driving more infrastructure deployments - your AWS and your Googles and your Metas of the world - but it's also your Cloudflare and those types of players who are driving these big infrastructure deployments. There are multiple places that we play with partners like that.

First off, we play obviously in the high-performance server space, so think of the general-purpose compute server. That is absolutely a place where we have a bunch of content, and we continue to work with them on things like CSS and IP in some cases.

Then beyond that, we also engage these customers in other aspects of their infrastructure, think DPUs and networking equipment. All their storage devices have Arm. There's actually Arm all over the data center in places that don't get talked about that often. Those are other avenues where we're engaging these customers. Sometimes, it's not directly. Often times we're engaging these customers through ASIC partners who take our IP, build silicon and then go feed into those devices.

The next segment is really enterprise. Enterprise is an area where historically we haven't done a whole lot in terms of general-purpose compute, and that's because of the dynamics around a channel and having an actual product that we could sell into that area. These customers are going to be much less likely to build silicon. In fact, none of them build silicon. What they do is look for silicon off-the-shelf and buy silicon.

Our presence within these enterprise players has historically been limited to where our silicon partners go off and engage. Think of networking, security, storage. Then there are all the appliance-type devices, where we are all over them. We partner with companies like Broadcom and Marvell and others to go service this chunk of the market.

Now obviously with AGI CPU there is an opportunity to go after some more of the higher-performance type devices, and I can talk about that in a second here.

Then finally, we've got what we call wireless and edge. This is primarily dominated by a lot of the big telco guys, so think like the Nokias and the Ericssons of the world, etc.

Similar sort of story with these guys. It's actually more of a hybrid. In the cloud world, we engage almost always directly with those customers. Sometimes through semiconductor partners, but very directly with them. In the enterprise, it's almost always semiconductor partners on the wireless and edge space. It's kind of a hybrid. Opportunity to work directly with them. Sometimes they are licensees of ours. Sometimes they build their own silicon, but they also partner with our semiconductor ecosystem partners. Again, in each of these there's a slightly different go-to-market motion going on, and AGI is driving up demand across all three of these segments.

When we started the business, what we originally called the infrastructure business, it was really about developing IP and providing that IP to customers. In some ways, even the idea of developing IP specifically for cloud is a relatively new concept. We really just launched our first Neoverse offering back in 2019.

Before that, we would hand customers a big document and a phone number and say go for it. Go build a process. You can imagine the barrier to entry associated with that. We were helping customers by building software or trying to get software ported, but all of the aspects of actually building the CPU and the system IP and so on was not something that we had spent a whole lot of time on.

That shifted when we launched Neoverse, and the impact of that shift was a dramatic reduction in the barrier to entry, a dramatic reduction in customer's ability to go off and build products. That's why in 2020, 2019 you saw AWS emerge.

Since then, we've continued to lower the barrier to entry. That's really what CSS was about. The move to CSS was about a couple of things. Yes, it helped us capture more value, but it also accelerated time-to-market for customers.

We saw that happen with CSS. Our first customer that we provided CSS to gave us a great quote, which some of you may have heard me say before: they told us that it saved them about 80 man-years worth of effort. Going from - because they had CSS as opposed to discrete IP - saved about 80 person-years worth of engineering work.

We had another customer who, from the time we handed them the CSS to the time they had production silicon back running Linux, including time through the fab, was less than 18 months. That's kind of unheard of in the semiconductor industry.

So it was really about accelerating time-to-market and helping those customers get there, and then obviously accelerating our own value capture.

The result of all this since 2019 has been pretty clear. Whether that be with the traditional hyperscalers or with others, it's really been this kind of hockey-stick-like ramp that I was showing earlier today.

At the end of the day, Arm Neoverse is really about how we deliver more growth, more cores, higher ASP, and greater volume.

This chart is super impressive, and I think one of the most impressive parts of this chart - which I don't think people quite grasp - is that from 2019 to really 2025, that's actually tremendous volume. And we're moving forward, but things are really starting to kick up.

People ask why, and this is the answer: we're just getting going.

The reality is that even the CSS story - some of these hyperscalers are literally just starting to deploy their products in earnest now. They've launched them, and they're really, really launching them. We have visibility now into what their plans are, and that's what drives our confidence level just in the base CSS business.

This is before I talk about AGI CPU. This is before I talk about things like 6G base station refresh cycles, or the impact on networking, or how the enterprise is going to react to agentic. This is before I even talk about that.

I've got customers - hyperscalers, the largest consumers of compute - that are just now taking their products into market, and that's driving a tremendous amount of upside for us. We see that kind of happening moving forward. It's super exciting.

By the way, as I said earlier, the whole goal of CSS and the whole goal of the IP was to accelerate and allow these customers to get to market faster, to build more products, etc. And you're seeing them now start to get to market.

If you look at it another way, it's actually really interesting, because you see exactly what I just described. You had initial IP, you had one or two customers that were able to take that, effectively productize it, and drive real volume. Then since we launched CSS, you see that ramp continue to move forward at a pretty amazing rate.

Now what's really interesting about this slide - both Drew and Chris talked about this idea of doubling and doubling, and we've got a similar dynamic going on here. I hadn't thought of it in that way, so I won't frame it that way, but structurally it's the same thing.

You've got multiple tailwinds happening here.

First is the share tailwind, which I just described. You've got these hyperscalers who have now deployed products or are now deploying products, and so that's accelerating and moving forward. Great.

We get more ASP per core because we've moved to CSS. That's also good. So you're capturing share and higher ASP - that's a good story.

Then you've got a growing TAM on top of all that. So not only are you capturing percentage share at each of these customers and capturing more ASP value at all these customers, but you've got increased unit share as well.

That's what's driving all of this revenue upside.

Now for me that's exciting. I hope you guys are excited too. But I've got to tell you what's even more exciting in my mind is kind of the story of the day, because you've got all these growth drivers in our base IP and CSS business, and then we have this massive market expansion that has now opened up to us.

That market expansion means that we have this opportunity to capture so much more value from those existing customers while also addressing a bunch of customers in new ways that we historically couldn't.

Remember that whole enterprise piece, or all of those tier-two hyperscalers, which even with CSS really can't go off and build their own silicon? Companies like Meta - these are all places where we weren't able to capture with the legacy IP business. So now we have SoCs to go after them, which I think is creating a tremendous opportunity for us.

The only thing I would say about this slide - and we talked about it a little bit earlier today, I think Rene mentioned it and I mentioned it as well - and I really want to drill this point home because it's so, so important, and candidly one of our biggest strategic advantages in my opinion, is that we'll give you any one of these.

If you're a hyperscaler and you're looking to build a GPU, sure, I'll give you IP. If you're a hyperscaler and you want to build a chiplet or a piece of silicon that is very tightly coupled to your accelerator and you need a CSS, I've got that for you too. You want to buy silicon off-the-shelf? Sure, I can give that to you.

Now you can use any one of these things across your infrastructure.

The reality is none of these players are only buying one piece of silicon. They're building complete systems, and now we have an offering and an option for them across that entire range that is very uniquely Arm and something that really only we can do.

That allows us a tremendous amount of leverage going forward, because it allows them a tremendous amount of leverage in terms of their software. I don't think most people quite capture the significance of that when they think about their overall infrastructure plans.

So this isn't just about market expansion. It's about making sure that we can go off and address and meet those customers wherever they are.

Let's talk about some of the specific customers we have for Arm AGI CPU and the diverse customer demand that we have, because this actually feeds back into that same point.

Let's talk about the AI data center for a minute. What you see up there are a bunch of customers who are looking at using it for the agentic use case - the one that we talked about earlier. Remember when they said agents never sleep? I love that line, because they don't. They keep going all the time.

Arm AGI CPU is great for that, but it's also great for that head-node capability. A lot of these guys are building accelerators, and guess where they're getting some of the IP for themselves? Not the core IP, but some of the connectivity IP - they're getting it from us.

So what can we do? We can provide AGI CPU, and then we can provide some of the interconnect technology or some of the other technology on their side to make that link, that connection, that much better.

That's part of the story that we think about with one of these guys - in fact, most all of these guys are our licensees as well, which is an important part of the story on the cloud side of the house.

What you see is actually a couple of different interesting use cases emerge.

If you look at SAP, that's a really interesting customer, because what SAP has done is they've actually moved their entire HANA database over to Arm in the cloud. They're loud and proud about the fact that they're an AWS Graviton customer, and they're going to continue being a great AWS Graviton customer. We helped them through that porting work as part of our IP and CSS business.

The reason why they're interested in AGI CPU is very simple: they're seeing all these benefits from Graviton in the cloud, they run a hybrid environment, and then they look at what they're getting in their data centers

because they've got some sovereign data requirements in places like Europe, and they can't necessarily get the same benefits there that they're getting in the cloud.

So all of a sudden they're like, "Hey, how do I get the same benefits that I'm getting with Graviton but get it in my own data center?" AGI CPU allows them to do that.

Guess what? You heard me say earlier - there are well over 10,000 customers using Arm in the cloud today. Those are all potentials for AGI CPU.

Cloudflare is another great example. They fall into a slightly different camp. I don't want to be derogatory and call them a tier-two hyperscaler, because that's not really what I mean. What I mean is they're not at sufficient scale to go off and be able to build their own silicon. They'd love the advantages that you get with moving to Arm, but they haven't had a product. They've been underserved. So how do they get those TCO benefits? Arm AGI CPU allows them to. That's why they're interested in Arm.

Then you look at F5 Networks, and I like these guys for a completely different reason, which is they build what I would effectively, lovingly, call an appliance. They build a bunch of these appliance boxes - networking, storage, and security. Typically these folks have a bunch of SKUs in their range. They've got high-end SKUs and low-end SKUs.

The low-end SKUs usually have to be pretty power-efficient, and many of them use Arm for that. We have a tremendous presence in that area.

At the higher-performance end, there was no off-the-shelf offering that would work. So what did they have to do? At the low-end of the range they had to use one code base because it was more efficient, and at the high-end of the range they had to use a different code base because of performance.

The reason these guys love Arm AGI CPU is that it allows them to streamline their entire code base. It becomes a very straightforward story. So they can go off and adopt the technology. Tremendous opportunities here, and that's kind of the way we think about the market when we go off and attack it.

When we look forward and we talk about our roadmap, the things that we're going to continue to lean into are those same things that have made Arm AGI CPU so exciting to customers already: best-in-class performance, best-in-class scalability, and best-in-class efficiency.

You heard me say it earlier and over and over again today, and those really are the things that we're going to lean into. We're going to make sure that we're leading in terms of interfaces. We're going to make sure that we're leading in terms of memory and IO. That includes looking at things like NVLink, next-generation PCIe, and different memory technologies to make sure we optimize for those three key attributes over and over and over again, because we are mission- focused on solving this problem for the industry.

Of course, there's a tremendous opportunity here. Jason's going to talk you through the numbers in quite a bit of detail. You saw the hockey stick. You know where the numbers are in terms of our royalty revenue. The opportunity for AGI CPU we think is enormous. We are incredibly excited about it, and I'll let Jason provide some more details in his session in a minute.

But underlying all of this - and the thing that I just want to make sure that I emphasize, like both Chris and Drew - is that our software ecosystem sitting on top of that Arm Neoverse compute platform, regardless of how you digest that Arm Neoverse compute platform, whether it be via IP or compute subsystem or AGI CPU, is the foundation to what has put us in such a position of growth and why this business is so exciting right now.

Thank you.