arm

Arm Everywhere event Investor Session full transcript

Please welcome Arm's Chief Marketing Officer, Ami Badani.

Hi everyone. My name is Ami Badani. I think I've met some of you, but probably not all of you. I'm Arm's Chief Marketing Officer. As you saw in the keynote, this is a huge, huge moment for the company, so thank you all for being here.

A little bit about myself: I've been at the company for two years. Prior to joining Arm, I was at NVIDIA, running the BlueField software business and the BlueField GPU software business. In my early years, I was actually an investment banker at both Goldman Sachs and J.P. Morgan, so I've had a bunch of different roles throughout my career.

We're going to spend the next several hours together, and I'll walk you through what that looks like. We're going to have each of the General Managers of the businesses come up and talk about the product strategy as it relates to our three businesses.

Chris Bergey will talk about the Edge AI Business Unit. Drew Henry will talk about the Physical AI Business Unit. Then Mohamed Awad will talk about the Cloud AI Business Unit. They will all talk about the product strategy.

Then we will have Jason Child, who's our Chief Financial Officer, come up and talk about the financials at a company level. At the end, we will do a 45-minute Q&A, where we will have all of the general managers on stage, including Rene and Jason. So please hold your questions until the very end.

Each of the business units, again, will go through their strategy, and then you can ask questions to each of them at the very end.

Also, I wanted to introduce Sharbani, who will go into the software section after all of the business units, because that's a very important piece of our business that we wanted everyone to understand.

So with that, I'll hand it off to Chris Bergey, who will talk about the Edge AI business.

Please welcome Arm's Executive Vice President of Edge AI, Chris Bergey.

[00:20:20] All right. Thank you all for being here. Obviously, an exciting day for Arm.

I was asked just to give a little bit of my background. I see some old faces here in the audience. For those who don't know me, I'm actually going to celebrate my 30th year in the

semiconductor industry this summer, after starting out of engineering school at AMD, then spending almost a decade at Broadcom, and actually joining Arm a little over six years ago. That was the first time I actually worked for an IP company versus a silicon company, and I did some startups in between there.

So for me, it's exciting to see Arm get into silicon, but I'm super excited about what I represent and what we're doing inside of the Edge Business Unit.

As you're probably aware, the Edge Business Unit actually represents a lot of the legacy Arm business. As large as it is, and as strong a footprint as we have, many of you may be surprised to see the amount of growth that we believe we're going to see over the next five years from a TAM perspective, and that's about a 40% increase in TAM.

Now, why is this happening? It's not a big surprise. AI. Essentially, what is happening is AI workloads are coming to all kinds of devices. Obviously, it's been pretty exciting to see the OpenClaw moment that Rene mentioned earlier, and this has been really fun for us because we've been talking about agentic AI for over a year now. It was really hard for people to necessarily connect and have that "aha" moment around why I would want a personal agent and why I would do these things.

Now you're starting to see these technologies come to market. But what you're seeing is that you start seeing the concepts, and that's great. But I think the other thing you see with many of these early agentic use cases, or even some of the prototyped agentic OS work being done on the development side, is that the concepts are great, but the performance is not there.

I can tell you that this has come through in spades for me. I actually just spent the last several weeks traveling between Mobile World Congress in Barcelona and Embedded World in Germany, and the number of customers that came to me and said, "Chris, you know what? We need to up our compute. We need to think about going to more advanced nodes than we were thinking about. I need to think about a more powerful CPU complex. I need to think about more memory bandwidth."

So, this is really what's going to drive much of this growth that we're expecting.

When you look at where the edge AI opportunity is for Arm and inside of my business unit, obviously mobile continues to be an incredible driver for us, and these devices continue to be essential to us. We believe that many of the agentic services that are going to start running on top of those devices are going to drive an incredible amount of additional silicon content.

One of the questions I get there is, "Wait a minute, Chris, what about the dynamics around pricing?" We're obviously seeing some of the tensions that are happening already today around memory and storage. But when you actually look at it economically, and you look at the cost of

many of those tokens in the cloud versus doing them locally, there are quite a few economic drivers, quite a few performance drivers around latency and real-time response, and then there are also a lot of requirements around privacy and where that data stays, and how do you protect some of those things, and where the agents have access to that data.

So, we see a huge amount of growth opportunity in mobile, and I'm going to talk more about that as we get into our CSS strategy.

The second thing is intelligent edge. Obviously, this is a huge category of devices, but what you see across these devices — whether they're consumer devices, enterprise devices, or industrial devices — is AI coming into those devices, and people looking for new use cases and new capabilities.

One of the analogies I like to use is touchscreens. Remember, the iPhone is almost 20 years old, the iPad is about 15 years old, and that was when we really started using touch. Now if you give a child a screen and it doesn't have touch, they basically give it back to you and say, "I don't like this. It's not the experience I know."

What you're really going to see in all these devices is that's how essential AI is going to be. The AI experience, the device intelligence, being able to sense things you want to do, being able to have conversations with it — all those kinds of activities. So we see this across the board, both around existing categories of devices and then new devices that you can see around XR platforms and the like.

Lastly, we see a lot of exciting things going on around personal AI computing, and how AI changes the way we have these compute paradigms in our homes and in our lives.

[00:26:00] The first thing to touch upon — and Rene obviously did a great job this morning talking about this — has been the journey from IP to CSS. This has been quite transformative for us.

The conversation that I would have with so many of our customers were: "Chris, we're not having the best performance in the market. If you could just help me get the best Geekbench score, if you could just help me get the best SPEC score, I could get more for my product, I could get more market share, I could do better, and I'd be willing to share some of that with you, Arm, because you're helping me be successful."

So, this was really the birth of CSS.

What we did for CSS inside of the Edge AI Business Unit is we looked at what it really takes to make a world-class product. Really, what you need to do is optimize all the way from the transistor, through the libraries, up through the RTL, up through the physical implementation, and then as well work very closely on the software stack.

This is what we have done with CSS. I think you heard Y.H. mention it in his video — the president of Samsung, Sammy — around the amount of work that we do with leading companies like Samsung, like TSMC, and the like, to make sure we are optimizing that full stack.

So what we do is we give customers not just the RTL and the IP — the GPU IP, leading CPU IP, system IP — but we actually are able to stitch that together, and then we also give them the recipe. We say, here you go: this product on the latest node, at 4GHz, gets you this score.

For some customers, that's amazing because they say, "You know what? I'm new at this. I'm just an OEM. I want to get into silicon. Give me that recipe, and I'm going to cut and paste that right into my chip."

Now we have more legacy customers that say, "Hey, we think we're really good at that too," and we say, "Great." But what we give them is a benchmark, because many times their engineers say, "We don't think 4 GHz is possible." We say, "Here it is." Then they get a chance to try to beat it and try to do 4.1, 4.2 GHz.

This has become a very virtuous cycle for us in really making sure that our partners are able to do very well in the markets that they participate in.

You can see this in our results. At this point in time, as of the base of the quarter that's ending this month, we have now achieved the point where 25% of our mobile royalties will be coming from CSS. But we're not stopping there. We actually have a tremendous amount of growth, as you can see from the slides, that we believe we're going to be able to pick up as we're really able to enable even more partners to do even better in their markets.

Obviously, this agentic drive for more performance across all price bands becomes very, very interesting to us.

But the opportunity goes beyond mobile.

A couple of segments that I like to talk about: one is kind of the emerging XR-type segment. I don't know about you guys, but obviously I like Meta on the data center side — it was great to see Santosh this morning — but I think their glasses are awesome, and I think I own three pairs at this point in time, so I definitely am a full-time user.

It gives you an idea of next-generation platforms, and all of those glasses are Arm-powered. In fact, the wristband shown there is actually a neural wristband that you can use — when they have the display glasses, you can start turning up and down the volume, doing those kinds of activities — and that also is Arm-powered, based on our Ethos NPU and MCU.

Again, we're really able to help enable these next-generation platforms that are starting to use AI-type features in all kinds of different power envelopes, because that is one of the most

difficult computing platforms right now because of battery life, sensitivity to weight on your face, temperature concerns, and all those kinds of things.

But beyond that type of platform, which to me is also driving a whole rebirth in wearables, we're also seeing a tremendous opportunity around personal AI computing.

One of the products I'd like to point out is another CSS product: the GB10 coming from NVIDIA. What you're seeing here now is this new category of devices where people are saying, "Hey, I want to run agentic AI." You can see what we're talking about from an OpenClaw point of view. You see great YouTube videos and examples as people are playing with these things. We very much believe that there's going to be a tremendous amount of computing platform opportunity here.

In fact, some of the products in this category are sold out — you can't get them.

What are the attributes of these products? Well, they have a tremendous amount of v9-powered Arm CPUs. Two, they've got great GPU subsystems, whether that comes from Arm or whether that comes from other partners' GPUs. And third, they have an amazing UMA memory system that's able to provide the bandwidth those AI platforms need.

So ,we believe this is going to drive a huge growth platform.

[00:31:45] I just mentioned I came back from China. It is OpenClaw crazy there right now around the number of platforms and what people are trying to play with, and the future types of products that they want to build.

What I think is also exciting is how this is translating into what you see happening in the PC industry. What I really see happening there is a really exciting transition happening. Obviously, Arm has made a big push around Windows on Arm. We've made some nice gains over the last several years. But what we're seeing now is something totally different.

At the premium level, you're seeing people thinking about AI workloads in the personal computing space. It could be running agents, but it's also about the next-generation creator. When I think about PC creating, it always was about publishing websites, those kinds of devices. But now, that's not a creator. A creator is creating videos — AI-generated videos. They're using Stable Diffusion to create new graphics capabilities. Those kinds of platforms, again, are a tremendous strength for Arm's partners, and we believe that Arm will be able to participate in a very meaningful way at that premium part of personal AI computing.

What's also happening, interestingly, is at the opposite end of the market, which is really focused on the more traditional client use cases, you're starting to see all of these efficient computing platforms come out, whether it's Google thinking about how they're going to merge

Android and Chrome together for Aluminium, or potentially Microsoft and how they're doing 365 Link, and then lastly you have Apple coming in with their neo platforms.

With the tightness around memory and storage, I really believe you're also going to see this more efficient category really start to eat up a significant amount of PC market share. That's a great story for Arm and our partners because they participate in many of those markets today. We have amazing power efficiency, we can get amazing memory efficiency, and not to mention our GPU footprint.

I think many of you know, but maybe it's a secret to many, that Arm is the highest-volume GPU shipper in the world. In fact, we've shipped over 12 billion chips that utilize our GPUs. We're also making a significant amount of AI investment in this space around our GPU platform, so expect to see more from us and our partners in that area.

So, when you think about the CSS opportunity in this space, this is very different than mobile. This is not a space that we have a significant amount of market share in today. This is really a growth opportunity for us. You're seeing the projections that we believe, and this will be a significant driver for us as we bring those CSS platforms into many of these other computing areas. So this is a key strategic area for us, and we're getting tremendous leverage from what we're doing already on the mobile and GPU side.

[00:35:22] Lastly, I want to touch upon just a little bit about AI software. Sharbani is going to talk more about the importance of software. Rene obviously touched upon it earlier today, but I wanted to just nail this point: v9 is absolutely the most secure and advanced AI CPU architecture out there.

Richard Grisenthwaite was up on the screen — he is our Chief Architect — and what he and the team have put together with our partners is the most advanced platform. SME2, our Scalable Matrix Extension 2, is now shipping in the leading handsets, both from iOS as well as Android. What you're going to see is those platforms proliferate.

At this point in time, we're approaching about 50% value share on v9 penetration. That's going to actually go up to 85% in just the next two years. So you're seeing v9 go everywhere, and AI is going to enable us to take that into other markets as well because of the AI footprint that we have here.

Again, you can see it's the leading OSes, the leading applications and tools, and then lastly the leading AI frameworks.

Many of you know that many of the really tough challenges around AI are around software frameworks. How do you make it easy for users to migrate their workloads, for them to reduce the quantization challenges, and for it just to be easy to run AI?

A year and a half ago, or almost two years ago, we introduced Kleidi AI at COMPUTEX, and it has been transformational for us because Kleidi AI is a library that allows developers to integrate into their platform such that when their AI workload lands on an Arm CPU, it is able to use the advanced instructions that are possible in that platform.

This is what has been integrated now in these frameworks that you see below. But I'll let Sharbani tell more about that to all of you later on today.

With that, I'm going to hand it over to Drew Henry.

[00:37:50] Please welcome Arm's Executive Vice President of Physical AI, Drew Henry.

What an amazing day. Days like this don't happen often in a company's history, and for me personally, it's just an exciting moment to be participating in it.

I've had an opportunity to participate in a lot of new technologies being brought into various different markets over my career. This one ranks at the top for me personally.

As many of you may know, I spent quite a bit of time in the early days at Silicon Graphics, actually building out compute platforms, where we had to build our own silicon on these little tiny wafers that were like this big. Then moving on from there to NVIDIA and taping out and launching a whole bunch of 500-square-millimeter die platforms for all the compute GPUs when I was doing what I was doing for Jensen.

So I've been doing stuff in this industry for a very, very long time, and it's been fun to participate in all those other kinds of programs. But this one in particular is a big one for us.

Rene walked through the history of where things started when we went private. I joined Arm shortly after that, and joined Arm specifically to help get started the infrastructure business that's now this Cloud AI business.

As a matter of fact, I remember Mohammed and I one day standing in front of a whiteboard trying to figure out how we help customers build stuff faster, and that's when we first came up with the early ideas that we called the virtual SoC, which has now become CSS.

But that journey over that time, from when we were starting that to now, was not just about figuring out how to get new products into the marketplace or start the businesses. It was also about transforming the company.

For the last number of years, I've been involved in taking our physical design group — a physical design group that would do very, very low-level silicon optimizations for the top foundries across the world — and we would work with them to help build and make sure that they had the most optimized silicon in the wafers that they were putting through the foundries. Then we

worked very closely with our partners to make sure that they could get up to the clock rates and speeds like Chris was talking about.

That group eventually transformed into the group that Steve Halter runs, which is our silicon group.

At the same time, I'm working on how it is that we take our entire software ecosystem and get our software ecosystem optimized, because Arm, if we're going to be participating in CSS and doing silicon and stuff like that, we also have to participate and help people optimize their silicon platforms. So that's what I've been helping with for quite some time.

Rene asked me just recently to start this physical design group for us, and I told him I was delighted to do so because I think that physical design is going to be the most interesting computing platform in the history of computing.

Let me get into that for a second.

[00:41:05] Physical design, or a physical AI, to kind of understand what I mean by that. Physical AI is where AI is embodied in a machine, where that machine then, AI embodied, where the machine then has to sense, decide, and act safely in the physical world to be able to take action.

I like to think of it this way: the key metric of this particular industry is the latency, the time between a photon hitting a sensor and an actuator actually firing. If you're traveling 65 miles an hour down the road and something is detected by a car that's operating autonomously, that moment from when that sensor detects something to when the braking system or steering system fires is the key attribute.

It's an incredibly complicated computing problem. It is, I think, going to become one of the most profound computing markets in our industry and in history.

To understand a bit about the TAM here compared to what Chris showed and compared to what Mo has talked about today, we think about the Physical AI TAM as being relatively small compared to some of these others — of course, about $25 billion a year or so, moving up to about $50 billion a year.

At $25 billion a year, it's really not hard to do the math on that. The Physical AI space today is principally automotive platforms beginning to move to more autonomous platforms. 100 million-ish transportation vehicles are sold each year, so you can kind of get a sense for that market.

Industrial robotics, which is kind of the robotics space that's really generating value today in the world, is only about 500,000 units a year. It's not a big opportunity today, and that will grow

over the next five years or so. It's growing as a result of content growing — compute content growing — inside both of these devices.

The compute content is going from traditional automotive platforms that you see today into much more autonomous platforms, that have started to show up with robotaxis and the like. But the units aren't increasing dramatically, the compute content is. You're seeing a large increase in total compute content, which is driving this doubling of the TAM for this particular time.

But I had to beg and plead with the team: please let me show something a little bit past 2031. This is where I think the hockey stick happens for this particular market.

I think this market will grow in 2031 and don't ask me what year Drew will this thing hit a TAM of $200 billion, and while I can't tell you exactly what year this thing is going but I absolutely believe to my core that it is going to happen. It's going to happen. This hockey stick is going to go up high.

You've heard this from others, and this is really as a result of physical AI being embodied in robotics, in robotics and robotics platforms, and humanoid robotics platforms in particular. And its amazing how fast this will continue to grow. But I believe this is what's going to happen.

Frankly, I think the $200 billion undercalls it. I think it will be a trillion-dollar TAM sometime in that timeframe. But you've got to get ready for it.

Of course, this is platforms that span across a whole bunch of different types of form factors. In the transportation space, it's everything from traditional cars — which are becoming more autonomous now — and eventually, of course, autonomous robotaxis or the likes that you're seeing. If you're just cruising around here in San Francisco and you see them now all over the place.

But of course, it also moves into autonomous trucking, autonomous heavy machinery, and the like.

I'm working on programs right now with customers where we're talking literally about how it is you deploy an entire fleet of heavy equipment into a construction project, and that heavy equipment actually goes in and does all the work that needs to be done autonomously, which would be quite remarkable.

But then of course it's moving into robotics. These robotics platforms will be everything from humanoid form factors for doing all whole other bunch or type of projects, but then surgical robotics, and things used in industrial applications, warehousing, logistics and the like, security platforms, medical delivery via drones, food delivery, and all this kind of stuff.

All these things are happening. This is happening. The key thing across all of it is that compute is key to it.

This is what we've been getting ready for.

Similarly to what Chris was talking about for these businesses, for the physical AI business similarly we've been preparing for this by getting our platforms ready for helping people become more vertically oriented.

For the longest time, Arm of course just provided the individual components of IP, and generation to generation to generation we improved the capabilities of those platforms, even now through to v9. But it's not sufficient to just be able to offer all this stuff as bags of capabilities. You need to be able to help people build stuff faster.

That was the key thing we saw early on in the infrastructure business: the vertical OEMs that wanted to go off and build Arm platforms similar to what AWS was doing they wanted to get to market faster. They wanted to get their own silicon into the market. How do you do that? You do that by actually enabling them to be able to do that themselves.

So our taking, so instead of saying, "Here's all this stuff, try to figure out how to build it yourself," we come in and say, "Listen, we know it's hard to do this stuff, so we're going to curate it for you. We're going to design it in the compute subsystem. We're going to make it something that you're going to be able to build much more quickly into your platforms because you're going vertical."

I think we're in a decade or more of more and more companies going vertical in the platforms that they're building, for a whole bunch of different reasons that we can talk about sometime.

This move from individual bits of IP into curated designs — where we've designed the subsystem itself — has been a big, big, big move for us, and one that's worked out quite successfully.

What it means for our business is a move from v8-type architectures into v9-type architectures which Chris described the benefit of that eloquently. But that has resulted – we put a ton of investment into that and hence a lot more value into it, that itself has helped in a doubling of the rate that we collected on those types of platforms.

Then when you move into CSS, it's the doubling of doubling. We're building on top of v9, and we then again increase the amount of value that we bring to our customers. Of course, in exchange for that value that we bring to them, we get value back, and that's increased royalty rates.

To understand what it means to go faster time to market, or to reduce the amount of labor necessary to build stuff, this is hundreds of millions of dollars of value to our customers. Hundreds of millions of dollars of value.

We've done the analysis on it. You can't show up to your customers and say, "I'm increasing the royalty rates," without also showing up and saying, "Here is my financial analysis that shows you what the value is that we're bringing to you when you get a product to market faster, when you actually reduce the amount of time it takes for you to build it."

Things like being able to provide a compute subsystem so that they can actually tape out a chip and take that to market without making any changes to that chip — that's a huge economic value to them. That's the success that we've had with doing these platforms, like our CSS platforms.

These CSS platforms are now designed and optimized in the physical design space. They are designed and optimized now to be able to support this world of physical design space. This is the work that we've been doing.

As the world transitions over time, you have to understand where we're coming from, and this is where there are significant investments that we're making.

The world today — the vast majority of vehicles on the planet are traditional automotive platforms — and Arm has been in this space for a very, very long time. We created the Physical AI business unit just a few months ago, but we've been involved in this business area and these market areas for quite some time — decades.

So for instance, you might be surprised to know that in the last 12 months, just into the physical AI space, we've shipped over 2 billion Arm devices into our ecosystem to support this space. That's because there is an awful lot of compute that goes into these platforms.

Traditional automotive spaces goes everywhere from the actuation systems that manage braking, acceleration, and the monitoring of a bunch of different devices inside the vehicle, all the way through to the central compute that does things like keeping you in lane, helping you monitor speeds, and those sorts of functions, to the more sophisticated platforms that are becoming more autonomous.

That's all moving now to new levels of compute inside these autonomous vehicles. The big things that are changing are the architectures in these platforms. These architectures are evolving to become much more complicated, and again, with that more complication is more value that we bring into it.

An autonomous vehicle has substantially more Arm value that's brought into them than even traditional automotive platforms. But because we invest so heavily in those traditional

automotive platforms and help them advance, all that technology into those platforms then moves into the autonomy platforms as well.

What's happened in these autonomous car platforms is you're seeing an increasing amount of compute being put into them. Increasing amount compute is built on a of Arm. Again, our CSS platforms are very enabling of this, and our latest platform is called the Zena CSS.

The big thing now is that as you move from this progression from autonomous – automotive platforms with ADAS capabilities into autonomous vehicle platforms — where the vehicle is moving itself — when the vehicle is moving itself, that's a big investment not just in the platform that moves the vehicle around, that manages the autonomy where decisions have to be made in milliseconds or faster, but you also have to improve the information that's displayed inside the vehicle.

Because if you're not driving, what are you doing? You want to know where you're going, you want to see mapping systems, you want displays inside these vehicles, and in addition entertainment that's going to be provided to you.

So there's going to be an awful lot of change that happens inside these vehicles that is perfectly set up for the way our business models are set.

[00:52:21] Of course, all of this moves into robotics. The same platforms that are going into autonomous vehicles are the same platforms that are going into robotics and empowering robotics.

The reason why the world and the industry is so fast-ready, so enabling and ready to move from autonomous vehicles into autonomous robotics — and why so many autonomous vehicle companies are also doing autonomous robotics investments and building those kinds of products — is because the compute is the same. It's a much more complicated actuation system.

As a matter of fact, that humanoid robotics platform is hands down the most complicated computing system ever. It is incredibly complicated. Controlling all the actuation systems, being able to assure that you're able to move a pinky appropriately at the same time that you're walking, at the same time as you might be getting instructions, at the same time that something is speaking to the robotics platform — it's an incredibly complicated platform.

So the computing has increased substantially, and therefore, as a result of that, our royalties could substantially increase. I think it's going to be a fast-growth product for us.

What's interesting, when you actually take a step back and look at it all, is that there are about four planes of computing that exist in these more autonomy platforms.

The perception-driven intelligence, which we are incredibly good at designing, has to run in robotics platforms and in automotive platforms. This is the platform that is responsible for how something moves around, how it makes autonomous decisions about how a car drives around the city, or how a humanoid robotics platform or some other type of robotics platform makes its way around a factory floor or a home.

That's a very, very, very complicated computing platform — one of the most difficult computing platforms you can build because of the real-time aspects in it. You've got to make decisions in milliseconds and nanoseconds.

Then there's a completely separate compute domain, which is the interaction-driven compute domain. That's where you're conversing. You're sitting in an autonomous car and you say, "Hey, wait, stop. I want to get a coffee right over there." The autonomous vehicle has to recognize exactly the context of that command, and then it has to communicate to the actual drive system that's managing the drive.

All of this stuff is incredibly complicated. Then it moves into how you manage the actuation system, and then into how the entire system is connected through the cloud. It's an incredibly complicated platform, and one in which we are really, really well situated.

Examples of that span the industry.

We've been working with the absolute pioneers in this space. Tesla, of course — I drive a Tesla. I drive 95% of my time now fully autonomously in my Tesla. We've been working with Tesla for a very long time.

We've seen announcements from Nuro, which is doing an autonomous platform to help companies that want autonomous vehicles go off and do that. They're building platforms using our technology.

Rivian just announced last December that they now have an AI platform that they've built for themselves — again, a company going vertical, and our ability to enable them to build that vertical platform themselves, being able to build the silicon platform they want using our technologies, v9 technologies, and our CSS technology, is what enables all this stuff.

Then robotics platforms like AGIBOT, who also build on top of the Arm platform today.

All of these companies are examples of real-life applications of people moving from v8 to v9, or moving from v9 into CSS platforms, and being able to use these systems that we provide.

[00:56:10] Interestingly, as Chris was saying, adoption in our space, the Physical AI space, adoption happens a little bit slower because it takes a little bit longer to put these platforms

into the marketplace because of the way that safety requirements are imposed on these platforms.

You want to make sure these platforms are safe. So you have to do a certification to enable that. So we move along little bit slower than some other industries in adoption of these types of technologies. But the adoption is absolutely happening. We're moving into v9 where v9 is part of the royalty collections and CSS, building on top of that doubling and doubling I referred to before, are going to again increase the growth of this particular industry for us.

This is all through the 2030s before the big, big growth happens as we move into the beginning of the next decade, which we are incredibly well positioned for.

That's why I tell people that we are just at the center of this incredibly large market opportunity that is staring right over the horizon for us. It'll be a $200 billion-per-year TAM. I think probably even more when it really hits its stride.

Like I said, I'm spending a lot of time in China looking at actuation systems and how those will drive down and become cheaper. This is how involved we are in the ecosystem.

This is going to be an incredibly fast-growing market for us when it hits that portion of the hockey stick, and we are really, really well positioned for it.

But we wouldn't be well positioned for it if we weren't making investments in software. That's my final point.

We do invest substantially in software ecosystems. As a matter of fact, we work very hard to migrate any platform on any other legacy architectures and make sure that they're optimized and moving into the Arm ecosystem.

As was described earlier, that's becoming incredibly easy for us to enable now for a whole bunch of different reasons. We have a very strong software ecosystem that we are able to rely on now that is incredibly mature, running all the platforms that we run. That means we are very very well situated for this growth into the physical AI space. Building on top of the investments we made over number of years in these platforms, from v8 to v9 to CSS and the like, the intimacy that we have with customers — customers going vertical, wanting to be able to build these kinds of platforms — we're bringing all the capabilities of our ecosystem together to enable this. But it does all require this very rich software ecosystem.

All right. I'm going to turn it over now to Mohammed, who's going to come back up and close on the opportunities now in cloud.

[00:59:08] Please welcome to the stage Arm's Executive Vice President of Cloud AI, Mohamed Awad.

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 cloudmore 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 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 look to 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.

[01:04:38] 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.

[01:07:29] 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 agentics. 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.

[01:11:13] 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.

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 I/O. 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.

[01:18:34] Please welcome Arm's Vice President of AI Services, Sharbani Roy.

Hi, I'm Sharbani, and I'm one of the over 2,100 teammates that we have here at Arm obsessing every day about software.

I feel like Drew and Chris and Mohamed kind of did my job for me already, but it's probably a good time to underscore: you need software to unlock the power of incredible hardware. You can't have hardware without great software, and that's why the breadth, depth, and value that we provide to our partners through our software ecosystem is our advantage.

Our ecosystem is massive. Rene already called out that we have 22-plus million developers building on Arm. Twenty-two million. I'm pretty sure that makes this the biggest software ecosystem there is.

Now when I think about this breadth, it covers our full platform. I wanted to pull all of these logos together that you've seen. We covered each of these kinds of individually, and this is just a representation of the various companies, the various partners that we work with. But it goes from cloud to edge. It goes from the phones that are in your pockets to the laptops that some of

you are working on right now to run amazing reports talking about all of the awesomeness that we are bringing forth here today.

It's helping to power the agents that you're going to use to help you write your next reports, to help you figure out what your next strategies are going to be. It's in the cars. It's in the robots. It's everywhere. This ecosystem is absolutely massive.

It's not just about it being the most ubiquitous compute platform. We also have incredible depth. We go up and down the stack. If you look, you can see we've got the best operating systems all the way up to your favorite applications, like Santosh talked about earlier today.

When it comes to depth, the depth of our software ecosystem is our durable advantage. We've been at this for a while. We've been doing this for 15 years. We have over 1,300 open-source and upstream projects that we work on, and we have over 50,000 partners — 50,000 companies that we collaborate with. Twenty-two million developers building on Arm, 50,000 companies — that's a lot of elbow grease. We've been working really hard, right Drew?

And like I said, we work up and down the stack. We start off at firmware, doing kernel-level optimizations. We're working up through open source, through all the operating systems, all the way up to the applications. We're incredibly strategic about how we think about where we need to optimize software.

That's a lot of software. That's a lot of territory to have to cover. So we have to think about how we're going to have the absolute best leverage and the best ROI on our investments in each of these projects.

That means a lot of times we're thinking about how we have optimizations at the operating-system level or, like Chris was talking about, our Kleidi optimizations. We want to be thinking about how we're not just working with partners, but also how we're putting things out into the ecosystem, helping our partners with their proprietary stacks, so the software optimizations that we provide just run seamlessly from cloud to edge and work behind the scenes most of the time like magic.

When I think about what we want to obsess over when we're trying to drive the depth of these optimizations, I'm going to tell you all the same thing I tell my team every day: we obsess over three KPIs — performance per watt, making it as easy as possible for developers to build on Arm, and making it as fast as possible for them to deliver innovation into the market through those applications that each of you love to use every day.

These advantages only compound with the AI economy, with our agentic future.

[01:23:07] How many of y'all are using agents today? Going to hope a lot more after this session as well.

I'm going to let you in on a hot tip: I like to think about agents like apprentices. Each of you are great artists in what you're trying to do, and an apprentice can only help you amplify and provide all of the additional support you need to have your next great innovation.

To do that, they have to do a lot of experimentation. They're running all the time. They've got to be running as fast as possible. That means we're going to have to swap in and out a lot of different types of software.

So when you send your agent to go and run off and solve something for you, you want it to be able to have good judgment and be able to quickly swap out different models or different cloud providers so that you're not just blowing the bank on your latest experiment, or if it's something that's absolutely mission critical.

The good news is, the agents are going to deliver the work to your fingertips — to your phone, to your laptop — because when you think about our ubiquitous platform, that's all running on Arm. The agents are going to be delivering the value where Arm is actually serving you the best.

Then the software ecosystem comes into play because our software has been completely optimized across the board from cloud to edge, so these agents can really easily go and pick and choose software that's going to work on those devices. They don't have to think about it. It's just going to work. It's going to run great on Arm.

These advantages only compound. So it's breadth, depth, and I want to talk a little bit more about the value.

We provide our partners with incredible value by helping them gain horizontal leverage across the Arm platform. I want to jump into an example with Meta. Seems appropriate — we heard a lot from Santosh, from Paul today.

Meta was facing a challenge, and we've been working with Meta for years, not just on the hardware side but also on the software side, and this is something that's near and dear to my heart as well.

You think about Meta: Meta is famous for having a single stack, a monorepo, where you have a developer who is pushing something new, building one of your favorite applications and a new feature, and whenever they push something out into the internal code base, everybody's able to go and pick that up. It's got to meet an incredibly high bar, but you have to have horizontal leverage within the company so anybody can pick up the best-in-class, the latest and greatest, because they're always at the cutting edge.

This means that when we're partnering with Meta on software, we have to be at our absolute A game.

Now they had a challenge because they wanted to unlock the value that we have from some of our new architecture and from our CPUs. Chris talked a little bit about Kleidi and our SME2 capabilities. Now that's incredible performance and incredible perf per watt that we're driving for.

[01:26:13] But when you think about how models are actually made, you're training massive models in data centers, right? You've got to use a ton of compute. You can have billions of parameters, and you've got to squish that down and put it onto a device. Again, how am I going to fit that in my pocket so I can go be checking my favorite update on Instagram or sending my family great videos on WhatsApp?

So, what we knew we needed to do was jump in and help them with optimizations through their AI frameworks.

Many of you may be familiar with PyTorch. It has been the fastest-growing framework. I have a little bit of bias — I used to work in frameworks at another company — so I love all frameworks, they're all great. But I can see the power of needing to take those AI frameworks, which can become quite large, and the models are larger, the frameworks themselves are large, and we needed to be able to help Meta optimize once and deploy many, to fit these models onto the devices that we're all obsessing over.

So, what we ended up doing is we worked with them through PyTorch. Arm actually has an active seat on the board. We help out with the Technical Advisory Board, and we help direct both PyTorch in open source as well as work directly with Meta.

Then we also provided incredible contributions to ExecuTorch, which some of you may have seen went 1.0 last October. That means, again, it's not only benefiting Meta and all of the developers inside Meta, but it's also benefiting open source altogether.

So now when you're going out and some of you might be playing around with PyTorch models, you don't have to go through a whole long path to deploy it on Arm. You can actually go through ExecuTorch, you can go faster, it can be easier for you, and you can still get the best perf per watt that you get on the Arm ubiquitous compute platform.

So that's a repeatable performance gain and efficiency for Meta, and that's also something where we've been able to take those lessons and extend them more broadly into our ecosystem.

Again, we have breadth — a massive, massive ecosystem. We have the depth, and we're strategic about how we help up and down the stack. We work directly with our partners, both on their proprietary stacks and to help them in open source, so that we can have these repeatable gains that compound and will compound for years to come.

All right. I think next Jason's going to come up and give us a little bit of an overview.

[01:28:52] Please welcome Arm's Executive Vice President and Chief Financial Officer, Jason Child.

Hello everyone. My name is Jason Child. I'm CFO. I think I've probably met all of you, and probably multiple times, so it's great to see you again.

Very exciting day. Exciting to walk you through some of the numbers. I imagine many of you have been waiting to see what it all adds up to.

What I'm going to do today is make sure that the business case and the financial model behind the Arm AGI CPU is very, very clear.

There are really three points to this.

First, demand from new and existing customers is allowing us to materially expand our opportunity through selling chips.

Second, our existing IP business continues to have strong underlying growth drivers, and the chip business is compounding, and not displacing, the IP business.

Third, when you put those together, the combined model has much larger revenue, profit, and EPS potential by FY31.

All right. So, introducing the first phase of Arm's market expansion.

Let me start with the framework. Customer demand for Arm-created chips and the size of the opportunity has led us to explore the chip market over the past three years. After exploring the market and our own capabilities, we're introducing, as Rene and the team have talked about, the Arm AGI CPU.

Furthermore, the Arm IP business is going from strength to strength. Royalty revenue growth and license revenue growth both have multiple long-term structural growth drivers, which we expect to continue for years to come.

Finally, the financial consequences by FY31, that the combined model is materially accretive to revenue, gross profit, operating profit, and EPS. Importantly, much of the investment is already in the business, so the additional chip gross profit has meaningful drop-through to earnings.

Arm's opportunity is huge, and we're growing into the largest market in history — one that is just getting bigger and bigger. It's over $500 billion today, and we think this grows to more than $1.5 trillion in FYE31.

Just to be clear, this is just semiconductor logic — CPUs and XPUs. No memory, no optical, just chips where you might find Arm technology either today or in the future.

Breaking that down a little bit more, Cloud AI includes cloud compute, both CPU and XPU, enterprise compute and networking, supercomputers, and so on. This is a $330 billion market today, growing at over 30% a year to around $1.2 trillion in FYE31.

Edge AI includes the chips that go into smartphones, consumer electronics, and IOT devices. This is a $180 billion market, which we expect to grow at around a 7% CAGR over the next five years to $250 billion.

Physical AI, which includes automotive applications and robotics, this is a $25 billion chips market today, doubling in size over the same period.

You've heard Drew talk about how the autonomous vehicles and robotics opportunity could be larger than both Cloud AI and Edge AI. However, our current view is that the inflection point likely happens after fiscal 2031. It could be earlier.

This adds up to over $1.5 trillion.

[01:32:43] Double-clicking on the Cloud AI market, we have greater than $100 billion of Cloud, AI and Enterprise data center silicon, and $55 billion for wired and wireless networking. The remaining $1 trillion includes data center accelerator chips. We will come back to that opportunity another day.

Today we are just focused on the CPUs, and there is $455 billion of non-accelerator CPU TAM that is an Arm sweet spot.

I want to direct particular attention to the $100 billion-plus in the data center CPU TAM, as that is the market we are addressing today with the Arm AGI CPU.

These numbers may be higher than you've seen before. However, we have visibility into our customers' roadmaps. We are confident that the demands that inference and agentic AI will put on CPU performance will drive healthy volume growth and ASP increases.

We're expanding our opportunity in three ways.

Firstly, with our IP and CSS offering, the only customers we can address are the large hyperscalers who want to build their own chips. Our maximum revenue is limited to a fraction of the chip value. Assuming a 100% market share at a 10% CSS royalty rate, this adds up to $2.4 billion.

Supplying a complete chip allows us to address the full value of this market at $24 billion.

Secondly, as only the largest cloud service providers we are building their own chips, we are expanding the opportunity by offering products that all data center companies can use — from the largest hyperscalers to neoclouds to telcos and enterprises.

Thirdly, over the next five years, we expect the size of the market to increase significantly, driven by the expansion of CPUs used for inference and agentic AI.

Together, this takes our total opportunity from $2.4 billion of possible royalty revenue this year to more than $100 billion of revenue in FYE31.

Not all of this will be captured by chips. We still will be offering our customers CPU IP and compute subsystems, and this will help us maximize our total revenue from the data center by allowing our customers to choose the right solution for them.

[01:35:16] This slide makes the economics, and thus our rationale for entering the market, more tangible with a simplified hypothetical example using an illustrative $1,000 chip price.

CPU IP has about a 5% royalty rate, so for every $1,000 of chip sales, Arm receives about $50 in royalty revenue. This is at 100% gross profit, so you get $50 of gross profit dollars.

Compute subsystems have about twice the royalty rate, so it generates about $100 in royalty revenue and gross profit. For or a chip.

A $1,000 chip revenue model delivers about $500 of gross profit dollars. So while IP and CSS are extremely attractive on gross profit margin, the chip model produces more gross profit dollars per chip. This could be an order of magnitude higher than the IP model.

That is why we are pursuing all three models.

It expands our market to include customers that were not interested in an IP model, gives our current customers choice, and for Arm it creates a much larger profit opportunity.

We're not going to force any of our existing customers to migrate to this new model. We are welcoming customers to stay on their current IP/CSS model should they choose. Should they decide to embrace the silicon model, however, this chart illustrates how that decision would be significantly positive for gross profit dollars.

All right, so how are we landing the new chip business?

This business is based on customer demand from multiple companies and across hyperscalers and large enterprises. These are companies who prefer to buy a chip from us over building one from our IP.

We believe that Arm is uniquely positioned to build a CPU for the data center. If you look at all the Arm-based CPU chips, most of the technology already comes from Arm.

Because we have such strong initial demand, we have been able to quickly turn customer interest into actual business. As you see, we already have multiple customers lined up, and we have line of sight to more than $1 billion in chip demand over the next two years. The vast majority will fall into FYE28.

Our biggest challenge is not finding customers who want our chips. It's actually memory shortages limiting our customers' ability to deploy our chips.

We expect material revenue from the Arm AGI CPU starting in FYE28, with an exponential ramp to around $15 billion in FYE31.

The first driver is increasing demand from new customers, naturally increasing chip volumes. We also expect increasing chip volumes, as well as rising complexity, to lift ASP significantly by FYE31.

Turning now to our IP business, starting with royalties.

Arm's royalty revenue has multiple secular growth drivers. The end markets to which our technology is being deployed are growing. We are gaining share. Our customers deploy more Arm-based chips. Increasing complexity is driving up core count, especially in the data center and high-end automotive chips, which leads to higher royalty per chip. And our most advanced technology commands a higher royalty rate.

Over the past five years, royalty revenue has grown at about a 14% CAGR. This has accelerated to over 20% in the past two years as Armv9 and CSS have started to ramp.

Looking forward, we expect that royalty revenue CAGR will be 20% over the next five years.

It might not reach 20% every year, as there will still be the occasional market downtick or inventory correction. Our licensing revenue overperformance in the past few years lends confidence to these future royalty growth rates.

We have discussed for a long time our efforts to add more value to our customers and to be compensated for that value. AI is a significant tailwind to our journey, as customers are pressed for both more compute power and faster time to market.

We all know we've essentially doubled the royalty rates from Armv8 to Armv9, and again to CSS. We could also charge a significantly higher royalty rate for each Armv9 generation, I am sorry a slightly higher royalty rate for each Armv9 generation. And a higher royalty rate for each CSS generation.

[01:39:47] The increase in complexity of the chips of today and tomorrow is also contributing to our royalty revenue growth.

In the data center, for example, AI — including agentic AI — is driving our customers to design an increasing number of cores into their chips. Tracking both customer increasing core count over recent years and our customers' planned increase in future core count, we can see the number of Arm cores per chip increasing by about 20% per year.

The significant annual increases in cores and rising price per core as customers take more Arm technology is a big part of our confidence in continued robust growth in the data center royalty business.

One of the questions we often get is around our visibility into the future revenue trajectory. I think this tells the story.

Many of the contracts that underpin our royalty revenue forecast are already signed, and the royalty rates are already agreed in contract. We have delivered the technology. Our customers are building the chips and, in many cases, are already shipping the chips in high volume.

Looking out over years 2027 to 2031, 70% of revenues we're forecasting to collect are already covered with royalty rates set in contract. Even by fiscal 2031, the contracted base is still around 60%.

The remaining 40% is almost all with existing customers who we are confident will want access to the next generation of Arm technology, typically at higher royalty rates, higher core counts, and higher volumes than they do today.

We expect continued Armv9 adoption across all edge devices, from smartphones to smart glasses to smartwatches — pretty much every device with a screen. We expect a further boost from CSS adoption, not just in premium smartphones, but across all personal AI computing platforms, including in the PC space.

The combination of higher royalty rates from next-generation Armv9 and next-generation CSS will deliver outsized royalty revenue growth.

You can also see that we have a very sizable 65% of our forecasted royalties based on rates that are already under contract through FYE31.

Royalty contract coverage tends to be lower in edge devices than for Cloud AI and Physical AI devices due to very fast design cycles in consumer electronics.

Turning now to royalty revenue in the Cloud AI business, I've touched on the drivers of our expectation in rapid cloud growth already: rapid growth in the market, ongoing share gains, rising numbers of cores per chip, and delivering greater value per core create a powerful compounding story.

Our confidence is bolstered by the contracts that cover 85% of our expected royalties over the next five years.

We expect our healthy royalty growth in Physical AI to continue as cars, particularly autonomous and ADAS, continue to adopt more sophisticated silicon for the digital cockpit and driver assistance. We also anticipate continued share gains in this sector.

Our confidence is very high given the long lead times in automotive. Ninety-five percent of our royalties are under contract through FYE31.

As Drew explained, we are very excited about our opportunity in robotics. Much of this opportunity lies beyond FYE31 and thus is not captured in these figures.

Over the next five years, we expect that Cloud AI will be our fastest-growing revenue driver, even without the contribution from AGI CPU chips. When you add in chip revenue, it will surpass Edge AI in FYE30 and become by far the majority of our revenue in five years' time.

Finally, to licensing as you know, this has been growing well ahead of our expectations. At the IPO, we said that we would grow low single digits, which we quickly lifted to mid to high single digits, and it has more recently been growing over 20% per year.

This has been driven by a combination of the AI cycle, more customers getting access to Arm technology, subscription licenses and compute subsystem agreements, and the expansion of our licensing and design service agreement with SoftBank.

This license growth is the basis for the royalty commitment that you saw on the prior slide. We think all these drivers will continue, with the SoftBank licensing growing around high single digits, and with the AI cycle continuing to provide the majority of growth through more demand for next-generation CPU IP and compute subsystems at higher royalty rates.

Of course, the strong license revenue growth should lead to higher royalty revenue growth in the years and decades to come.

[01:44:58] As I mentioned right at the start, the Arm AGI CPU business, the royalty streams, and the licensing revenue all compound on top of each other. The chip business is targeting customers who either don't have the internal resources or don't have the desire to develop their own chips.

We do not expect the chip business to displace the IP business. Though some customers may ultimately choose to switch, it is accretive to earnings power as we previously discussed.

With around $15 billion of revenue from the chip business expected in FYE31, and another $10

billion of IP revenue, we are forecasting very strong revenue growth from the combined business over the next few years.

The good news is that we have already done a significant part of the heavy lifting when it comes to hiring the engineers needed to hit our plan.

If you've been following our financials for the past few years, you will know that we've already ramped R&D to support our product roadmaps.

Increasing R&D, combined with good execution, creates a virtuous cycle of new products driving revenue growth. From here, we are forecasting mid-teens OpEx growth through FYE31. Most of the incremental spending is our R&D investment in new technologies.

We expect our revenue by FYE31 to have grown by more than 2.5 times faster than our non-GAAP total costs. So as revenue and gross profit scale, particularly in chips, much of that incremental gross profit can drop through. That is the operating leverage in our model.

Our focus here is on the long-term earnings power of the company.

Before we get there, yes, we recognize that you have interest in the near term as well, and so we are affirming the Q4 guidance that we issued in February.

Back to FYE31: we see two meaningful profit engines.

First, we expect the IP business to reach about $10 billion of revenue, achieve a 99% gross profit margin, and deliver over 65% non-GAAP operating margin. We are today increasing our operating margin target by 500 basis points from our previous long-term target of 60%.

Second, we expect the Arm AGI CPU business to reach about $15 billion of revenue, with a gross margin of at least 50% and a non-GAAP operating margin of over 30%.

Putting those together, we have a consolidated business with $25 billion of revenue, industry-leading blended gross profit and operating margins, and more than $9 of non-GAAP EPS power in FYE31.

This is not a story of choosing between IP and chips. It's a story of combining a very high-margin IP model with a large, fast-growing, and accretive chip business.

Let me close on the three points that I started with.

First, customer demand is allowing us to materially expand our opportunity through selling chips. We already have line of sight to over $1 billion of demand from some of the companies that you met today, and are forecasting $15 billion of incremental revenue.

Second, our existing IP business continues to have strong underlying growth drivers, with the chip business compounding and not cannibalizing the IP business. We expect this to deliver around $10 billion of revenue in FYE31.

Third, by FYE31 the combined model is significantly accretive to revenue, gross profit dollars, and operating profit dollars, with more than $9 of EPS power. Because much of that investment is already in the base, the incremental economics are very attractive.

All right. With that, we're now going to conclude my presentation, and we're going to make a quick transition to start the Q&A session.

[01:54:58] Q&A session

arm