

## Nasdaq Investor Conference

### Company Participants

- Colette Kress, Executive Vice President & Chief Financial Officer

### Other Participants

- Mark Lipacis, Analyst, Jefferies

### Presentation

**Mark Lipacis** {BIO 2380059 <GO>}

Okay. Great. So do we have Colette on? Colette, can you hear me?

**Colette Kress** {BIO 18297352 <GO>}

Okay. How are you?

**Mark Lipacis** {BIO 2380059 <GO>}

Hi, Colette. How are you?

**Colette Kress** {BIO 18297352 <GO>}

I'm good. I'm good.

**Mark Lipacis** {BIO 2380059 <GO>}

Well, welcome, and so welcome to London. We are in the Jefferies room. We have nearly a standing-room-only audience here, and also some would like to welcome you. I would like to welcome, everybody, who came and stayed for the NVIDIA fireside chat and everybody who is dialing on online. So, thanks a lot, Colette, for making time for us. And I know that there was travel conflicts that prevented you from coming, but we're very grateful for you to join us at 8:30 in the morning, West Coast Time. So thank you for that.

**Colette Kress** {BIO 18297352 <GO>}

Thank you, Mark, for arranging this. It was a good solution -- conflicts, it's so difficult to spend that much time traveling. There's a lot to do, but I'm really pleased that we found a solution here.

## Questions And Answers

### Q - Mark Lipacis {BIO 2380059 <GO>}

Well, we found -- we understood that you were very busy processing orders from your customers so we completely understand. So Colette, I'll just go -- I'll just introduce you and then I'll just start with the Q&A.

And so Colette is the -- as many of you already know, she is the EVP and Chief Financial Officer of NVIDIA. She joined the company in September 2013, so coming up on the 10 year anniversary right around the corner. She previously served for three years as Senior Vice President and CFO at Cisco's Business Technology and Operations, Finance organization. And previously, she spent 13 years at Microsoft, including four years as CFO of the Server and Tools division. Prior to that, she served at Texas Instruments in a variety of financial positions. So looking at her resume, she has a very good track record for picking out the big and dominant players in each of those respective industries. So we'll be eagerly watching if you make another move.

So Colette, so I would like to start off asking you about a big picture question. And so if the outlook for you and Intel and AMD is accurate then this second quarter will be the first quarter in the history of the data center where NVIDIA's data center revenues will be greater than Intel and AMD combined. And I would argue that is proof of a tectonic shift in computing. And what are your reflections on that? What should investors read from that? And what are the key elements of NVIDIA's strategy that brought you to that position?

### A - Colette Kress {BIO 18297352 <GO>}

Great question, Mark, to start out with. But I'm going to do it real quick. As a reminder, this presentation contains forward-looking statements, and investors are advised to read our reports filed with the SEC for information related to risks and uncertainties facing our business.

So when we look at what we're seeing today in our data center business, it is not something that arose in terms of a short amount of period of time. We have been working on our accelerated computing platform, and all of our different solutions, our systems, our products. I'm working with the ecosystem for more than 15 years. And some of the work that we are seeing now is really starting to blossom.

Or another way to look at it is, it's an important -- inflection point. And it's an inflection point if you think about two different key pieces of that. The first is thinking about the importance of accelerated computing and how many folks around the globe have really understood the importance of moving to accelerated computing for better performance, for the next generation of computing and to focus in terms of on the sustainability that is going to be truly essential in all of the data centers that we see.

We have been working this for many years. Accelerated computing is not the same as just calling about accelerators, but it's really looking at the full data center as a whole and the decoupling of all of the different parts and to focus on how they can speed that together. Now keep in mind, the end of Moore's law really triggered more fuel towards accelerated computing, and that's what we're seeing today as people understanding that the \$1 trillion worth of an installed-base of CPU-only types of servers is in question in terms of, is that something that folks will want to refuel and readd to their infrastructure or is this going to be the time where they are moving to accelerated computing. The performance improvement that you could get right now with the CPU-only server is not high. That's a very cost-intensive move or not on a continuous improvement in performance. So we see accelerated computing as being a driver today.

That's not the only driver that we see. The other driver that we see is the inflection point that we're seeing in AI and more importantly, over the last six months, what we've seen in terms of Generative AI. We've talked for quite some time about the different types of workloads that are important right now or different types of solutions for AI, focusing on recommender systems, recommender engines also large language models. Generative AI really takes that large language model and now provides more offerings in terms of how for generative AI that can be front and center.

Another way of saying that is ChatGPT over the holidays, really opened up the eyes for CEOs around the world and essentially consumers or anybody in the enterprise of looking at solutions that they know they could use for assisting their companies in monetization or just improving efficiency by using and moving to Generative AI. There is a lot of different solutions. There is a lot of different models will likely be built just visualizing what they saw with ChatGPT.

So our guidance for Q2 is really building upon years and years of working of the industry on accelerated computing and solutions such as AI. And, we continue to see demand and demand visibility for the full year as well that we believe will sustain as we finish Q2.

## **Q - Mark Lipacis {BIO 2380059 <GO>}**

Great. And we think about on the demand-side, you talked about the generative AI and then I think one thing that also you guys have mentioned in the past is the power of the ecosystem. And I've heard Jensen observed before that the PC era was the only computing era that was a horizontal computing era where different companies made the chips and the software and the box, but every other era was that the ecosystem was a vertically-integrated from a single company, IBM and mainframes, digital, equipment corporation and minicomputers, Nokia in feature phones, Apple in smartphones.

So maybe if you could just share with us your view, like, what are the key parts of the ecosystem -- the software ecosystem that you're offering that a chip company may not have offered in the PC era?

---

**A - Colette Kress** {BIO 18297352 <GO>}

That is correct, Mark. We're not a chip company. Our whole position from the data center perspective is thinking about the full ecosystem and a full end-to-end stack, not just in terms of the time of a GPU is accelerating work, but the time that information or data enters into the data center. Or another way of looking at it is our ecosystem is focused on data center computing as a whole end-to-end. What does that mean? What does that incorporate? Surely, our work in terms of the GPU and that being some of the base work that we have done for more than 20 years, 30 years, but also focusing on a development platform.

So, way back when 15 years ago, the decision to make and put CUDA available on every single GPU we sold was key. And it was a key development platform that folks could take the GPU and determine what is a great new use-case, researchers, folks leaving higher education, all learning about CUDA and using that and continuing to build on that. We have many, many different generations of CUDA, but that has been some of the underlying pinning of why we have been able to scale is just that one decision more than 15 years ago. But there has been a lot more things that we have done with the ecosystem.

We take an approach of being agnostic to any of the different parts within the data center. A data center is filled from CPUs to storage, to security devices, new focus in terms of on a DPU, data processing unit, and much in terms of memory as well. Our focus is thinking about data as it enters and the speed of it transacting throughout the accelerated computing workload as a whole and building upon that with interconnect, our own NVLink connection, CPU to GPU, GPU to GPU continues to find the most efficient ways to focus on a very, very growing market such as AI. We're still in the early days of AI, although maybe here at an inflection point, there is a lot of work in terms of to scale-up or scale-out in terms of that data center.

But more importantly, outside of just CUDA is the stacks of additional software from SDKs to full libraries, helping industry by industry, help move to accelerated computing and help move to AI as we better understand the applications that they use within that industry and realigning it to GPU use versus just the standard CPU.

A lot of this is work with our engineers working end-to-end with our customers. We have, of course, a very large hardware engineering team solely focused in terms of not only our GPU, but also our networking through our acquisition that we did with Mellanox. These are important pieces. But let's not forget that we have engineers almost equal size in terms of software engineers that are focused in terms of building out so much of the work that we've done over this time.

**Q - Mark Lipacis** {BIO 2380059 <GO>}

If you compare the ecosystem that you guys have built to like what we may have seen in the PC era or like with Apple, I think there's a debate about whether or not you have an open ecosystem that get levered, you get leverage from by embracing developers versus kind of a closed one. And I think if you think about Macintosh back in the day that was maybe more of a closed ecosystem that limited their -- its

ability to scale. How would you characterize the software ecosystem that you guys have architected?

**A - Colette Kress** {BIO 18297352 <GO>}

Our software ecosystem has been the work with the industry players in providing us tremendous insight in terms of what is needed and how to rewire a lot of that work as you move to accelerated computing. But what it does not do is say, we leave it open to anybody to start working on that because honestly, it's not where they want to spend the time. When you've got a fast-growing market, as we do in accelerated computing and AI, we see folks really wanting to work on solutions.

Well, NVIDIA is working stitching that together and making sure that no matter what you are working on and no matter how changes that you make in terms of the type of GPU moving architecture, architecture, we will be backwards compatible and forwards compatible. So probably as long as you can imagine, more than 10 years working back and more than 10 years working forward.

Why is that an important step that we have taken? You can count on that continued software update that NVIDIA is doing and keeping it most current. And this allows any model to be opened and be used no matter which architecture you're on. That keeps us focused in terms of our software and consistency of our software across all of our different platforms. So it's a balance. It's a balance that says, yes, we absolutely need the input from the industry helping us in terms of what we need to do with the software. But you're right, we are on the hook to complete the software and make it available and easily available for any part of the ecosystem.

**Q - Mark Lipacis** {BIO 2380059 <GO>}

Great. I want to shift to some more near-term questions that we get. You guys gave guidance for the July quarter, which beat everybody's expectations and 50% sequential growth. And I think there is a concern from investors that this is kind of a one-time spike that will come back down. I know you guys only guide one quarter at a time. But what do you say to investors who have a concern that it's a one-time spike? And can you just talk in general terms, what is the visibility typically like with the data center and hyperscale companies?

**A - Colette Kress** {BIO 18297352 <GO>}

Yeah. What we have seen is certainly an astounding amount of interest worldwide globally from many different types of customer sets. That means our CSPs. That means our consumer Internet companies. That means our enterprises, our high-performance computing customers and supercomputing. All of these eras are important parts in terms of what we're seeing in terms of interest.

And so our visibility has improved. Our visibility is about companies helping us understand when they are ready to accept compute, where they want that compute to stand within their data centers and how we can help them even with the networking and compute together. So our visibility has improved. We have better visibility than what we've seen before.

And our ability to focus right now on procuring the supply. As we indicated in our earnings, we have procured the supply to the demand that's been put in front of us and that visibility that we see. So do we believe we're at a different stage than where we've been in the last several years in terms of the growth of AI? Surely, we do. And we do believe generative AI is something that we will fuel for several years for sure.

**Q - Mark Lipacis** {BIO 2380059 <GO>}

Got you. And I would say another -- and I know that you have fielded this question probably 1,000 times since you reported. There's kind of this view that all of your growth is driven by ASPs and maybe not as much by units. And how should investors think about the ASP growth versus the unit growth in the quarter past and the quarter outlook?

**A - Colette Kress** {BIO 18297352 <GO>}

No, we truly are seeing demand and need across such a wide group of folks. And that's not based on sure we are selling a brand new architecture and ramping a new architecture, a Hopper architecture. We are still shipping our existing prior architecture (inaudible). But no, this is not about ASPs. This really is about just the growth that we're seeing in focusing on AI and accelerated computing. I believe more of it is just about sheer volume of companies that are really interested in taking this next step and really leveraging generative AI for all the work that they do. That's really what it's about.

**Q - Mark Lipacis** {BIO 2380059 <GO>}

And can you give us a sense of -- I know that you're in a -- you're reramping a new product, the Hopper product. Your previous generation was Ampere. Can you give us a sense like where are you in the transition period? Do you think -- do you sell the Ampere product for a long period of time? Is it a long transition? Any insight on that, I think people would appreciate.

**A - Colette Kress** {BIO 18297352 <GO>}

Yes. Our Hopper architecture started hitting the market at approximately Q3 of last year or Q3 of our fiscal year last year, and we started ramping in Q4. Many of our cloud service providers were very interested in making it available in the cloud, setting that up and doing all the qualification for future H100 types of purchases. So that's where we began.

But it's very, very common for us, architecture to architecture to keep selling the existing architecture. Why? It was already pre-qualified. We have folks that are extending it to existing clusters that they have, and that becomes also just as great of an architecture because it is probably the second best in the world for them to continue working on Ampere. So it is likely that this dual type of architecture will continue. We'll think about this as when does it start to slow down, unsure at this time. But yes, we will still sell in both of those architectures.

**Q - Mark Lipacis** {BIO 2380059 <GO>}

Can you talk about your inventory strategy? I think normally, in semiconductor is when you see companies build up inventories, you get a little concerned that perhaps they mismanaged their strategic planning process, right? Now you built up inventories a lot in front of this previous quarter. So it looked like high inventories were actually a good signal.

What are you trying to do with your inventories? Do you want to -- they came down a little bit? They came down a bit? Do you want to take them down? Or do you want to take them up? Maybe talk about how your customers are -- what their kind of observations are on your inventories?

**A - Colette Kress** {BIO 18297352 <GO>}

Now let's look at it in the perspective that there is one piece of our supply chain which is looking at our inventory, our inventory on hand. And so this is always going to be in a staged approach. And remember, there's a large mix of what we have in terms of on our inventory because we have actually several other businesses, our automotive business, our graphics businesses of gaming and professional virtualization and data center.

So each of those different areas have different lead times, lead times in terms of when we are procuring, when we are accepting inventory. But again, a lot of it still has to be done to complete systems. So we may be in a staged approach until we actually finish the full system and ship the system.

But outside of just looking at our inventories also thinking about our purchase commitments. If you read in terms of our purchase claims that we also disclosed, that's where we are also ordering but we have not necessarily received it. Because remember, we are not the contract manufacturer of a lot of what we're doing. So we are procuring through contract manufacturing and then, of course, even at the wafers and/or from the memory and other different types of pieces.

So our overall strategy is to be as diligent as possible in terms of looking at our forward-looking demand and being able to back into what kind of lead times we have, so we can meet the overall customers' expectations in terms of when they want to receive that.

Right now, given data center is now a significant part of our business and longer lead times, you will see a different [ph] in terms of days of inventory and/or purchase commitments just due to the complexity of what we have to build as well as those lead times. For example, when you think about the Hopper architecture that is an architecture that has something more than 30,000 different components to create one of those systems. As much as we love the wafers, the chips part of it, that's just one part of our journey to complete that great system.

So our position right now is we surely are watching both the inventory and the purchase commitments together. But it's all about getting ready and understanding the demand that we have in front of us. That's really where our strategy lies.

**Q - Mark Lipacis** {BIO 2380059 <GO>}

Great. I just want to -- I'll come back to the kind of the ecosystem kind of strategy here. Software is part of that. You have previously disclosed that the software is a growing part of your revenue base, although it is quite small. How do you see the software growing? Do you think it grows faster? And do you ever think about -- what are the ways that you recognize kind of non-hardware revenues through software or through the ecosystem approach that you've taken?

**A - Colette Kress** {BIO 18297352 <GO>}

Correct. We've taken an approach of selling software separately. As you know, many of our systems right now incorporate a significant amount of software that is not really sold separately, but it is included in the full price of some of the systems that we create. But there's another opportunity for us to create software to help many of our enterprises and many different industries move towards AI. So let's kind of talk about kind of the three largest pieces of it.

The first one is NVIDIA AIE. NVIDIA AIE is essentially a full package of the essentials of an operating system to do overall AI. Important for the enterprises who want to get started that can really leverage this for keeping everything straight within their data center, keeping a close watch of all of their compute as they've got a full stack to both begin building what they need in terms of a new model and/or tracking in terms of the progress and how efficient the data center is running. All this is available. Enterprises want us. They want to make sure that we are helping and providing software that keeps their risk minimized and they can count on us completing that software. That's one piece of that.

The second piece is Omniverse. Omniverse is a platform solution software. This is allowing the future of the metaverses in many different ways, ability to create digital twins, the ability to create a 3D environment in the Internet. All of these are important pieces and how you can work with both individuals or teams within the enterprise as they look for Omniverse's solution.

And then lastly important one is our automotive software. Our automotive software sits in addition to the hardware that would be inside of the car, fueling ADAS or fueling Level 2, Level 3 and beyond type of software. That's with many of our partners, particularly our Daimler agreement as well as our JLR agreement. That will again be an important ramping as we are sharing the software revenue between us and those partners. So we're very excited for that piece as we do believe that our automotive pipeline going forward is about \$14 billion, and a very good portion of that is focused on the software revenue that we will have with these two partners and likely more going forward.

So right now, our software revenue, we've indicated is in the hundreds of millions of dollars and we do believe it's an important piece with the infrastructure purchases that software will likely be hand-in-hand for so many of the things that we see going forward.



**Q - Mark Lipacis** {BIO 2380059 <GO>}

Great. And we're running out of time, if I may ask one last question. You talked about -- if you just talk to your manufacturing strategy. Jensen has made the argument in the past that you guys like to have a diversified strategy. You use TSMC, you use Samsung as important suppliers. And at COMPUTEX, I believe he said that he's interested in exploring sourcing chips from Intel. Could you just play back like that announcement? Like what was the message there? What should investors think about that?

**A - Colette Kress** {BIO 18297352 <GO>}

Absolutely. Intel from a foundry perspective is very interesting. It's very interesting to many companies, and certainly, it's interesting to us. We love the redundancy that allows us to have multiple fab providers. We're probably one of the only ones that has a very strong two fab providers currently with TSMC and Samsung. Adding a third, looking for US also bound [ph] type of fabs would be great for us as well. So our work continues.

And a lot of that is really about the partnership. It's about the partnership that we have with TSMC, a partnership that we have with Samsung. Before you think about the machinery, it is really about the people, the service, and getting to know how each of us are going to operate. So we know that can be done. We've done it for 30 years and have continued those partnerships, and we'd love to see if we could add something like Intel as well.

**Q - Mark Lipacis** {BIO 2380059 <GO>}

Great. Well, we have run out of time. Colette, thank you so much for joining us today. Really appreciate you getting up early and spending some time with us. Your comments are very well received. And hopefully, maybe we can see you back here next year in-person.

**A - Colette Kress** {BIO 18297352 <GO>}

Absolutely. Absolutely. That would be great. Okay. Thanks so much for arranging. Appreciate it. Take care.

**Q - Mark Lipacis** {BIO 2380059 <GO>}

Thank you, Colette. And thanks for everybody for joining, and thanks for everybody online who dialed in.

*This transcript may not be 100 percent accurate and may contain misspellings and other inaccuracies. This transcript is provided "as is", without express or implied warranties of any kind. Bloomberg retains all rights to this transcript and provides it solely for your personal, non-commercial use. Bloomberg, its suppliers and third-party agents shall have no liability for errors in this transcript or for lost profits, losses, or direct, indirect, incidental, consequential, special or punitive damages in*

---

*connection with the furnishing, performance or use of such transcript. Neither the information nor any opinion expressed in this transcript constitutes a solicitation of the purchase or sale of securities or commodities. Any opinion expressed in the transcript does not necessarily reflect the views of Bloomberg LP. © COPYRIGHT 2024, BLOOMBERG LP. All rights reserved. Any reproduction, redistribution or retransmission is expressly prohibited.*