

Citi 2020 Global Technology Virtual Conference

Company Participants

- Atif Malik, Analyst
- Colette Kress, Executive VP & CFO
- Stewart Stecker, Director, Investor Relations

Presentation

Atif Malik {BIO 7312618 <GO>}

Welcome, everyone. My name is Atif Malik. I cover U.S. semiconductors and equipment stocks here at Citi.

It's my pleasure to welcome Colette Kress, CFO of NVIDIA. Colette is always a big draw at our conference with a keynote.

So it's good to be doing this virtually today.

The format of our discussion is a fireside chat. I will go with my questions first. If you have any questions for Colette or NVIDIA, please e-mail it to me. That's atif.malik@citi.com.

I believe Stewart is going to read disclosures first and then we'll start the fireside chat.

Stewart Stecker {BIO 17308457 <GO>}

Great. Thank you, Atif.

During this call, we may make forward-looking statements based on current expectations. These are subject to a number of significant risks and uncertainties, and our actual results may differ materially.

For a discussion of factors that could affect our future financial results and business, please refer to our most recent forms 10-K and 10-Q and the reports that we may file on Form 8-K with the Securities and Exchange Commission.

All our statements are made as of today September 9, 2020, based on information currently available to us.

Except as required by law, we assume no obligation to update any such statements.

So with that, I'll turn it back over to you, Atif.

Questions And Answers

A - Atif Malik {BIO 7312618 <GO>}

Thank you. Colette, I'm going to start with the gaming segment first, though it is a bit smaller than your data center sales at 43% in the July quarter.

That's because we're coming off a very exciting period, a very exciting launch on new gaming cards.

So the recent RTX 30 Ampere series launch surprised investors on both performance uplift of 2x over the Turing and pricing, which is similar to RTX.

Can you talk about why the team is so excited about RTX 30, the new apps like Reflex, Machinima, Broadcast and the upgrade opportunity within your 200 million installed base of gamers?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So thanks for the question. Thanks for hosting us here at your conference.

We look forward to this every year.

So let's start first with late last month from the kitchen of our CEO, we launched our Ampere architecture for gaming. This is our GeForce RTX 30 series.

So the GeForce RTX 30 series is our biggest ever generational leap that we've always had in gaming. It's up to 2x the performance of Turing, our last architecture, and 1.9x power efficient versus our last architecture.

It comes with a complete lineup of both architectural products for gaming, but it also sets us into the second generation of our NVIDIA RTX. Remember, NVIDIA RTX or ray tracing was first of its kind and setting the stage for the next generation of graphics of the future that we set out with the overall Turing architecture.

What is also key, which you highlighted, is with the overall 2x increase in the overall performance, we also have the same overall band of price points with the launch of our products that we have.

There's also some key features to keep in mind with this.

We are on the second generation of the RT cores. This is about 2x the throughput of the previous generation of ray tracing cores.

We're also on the third generation of our Tensor cores, with 2x the throughput from the previous generations.

There's also the RTX IO, which is about 100x overall performance increase with the HDDs and additional storage APIs that we've incorporated. This comes with the fastest graphics memory, the GDDR6X, so to improve the overall data speeds to close to one terabyte per system of memory bandwidth.

So there's a lot of great things just overall in the core of what we've provided with overall GeForce.

But you highlighted a bunch of other things that were very key in terms of our launch.

Other things that we show are support of both the ecosystem and our work in terms of the software for the ecosystem, to not only improve the overall development of games, but also to improve the experience of the gamers as they're overall playing.

One of them was NVIDIA Reflex. NVIDIA Reflex, as it may sound, is a suite of technologies that allows you to optimize and measure the system's latency. This is very key and will be a part of some of our important games that are out there, Apex Legends, Call of Duty: Warzone and even Fortnite, that reduces the latency by up to 50%.

Omniverse Machinima is also a very key app that is built on the Omniverse 3D workflow collaboration platform. This allows both cinematic production using a path trace viewer tool and engine design for physical accuracy. A very key thing, as you'll see, going forward, the importance of the live collaboration inside of games as well as other workloads.

Additionally, we can talk about what we're seeing in terms of NVIDIA Broadcast. NVIDIA Broadcast is understanding not only are people wanting to play games and/or watch games, but there's a very big universe out there that is working on broadcasting how they are working within the overall game. This is a universal plug-in that enhances the quality of the microphones and the webcams with RTX accelerated AI functions.

So a lot of different things that we've launched with the overall architecture of Ampere for gaming, not just focused on the overall cards, but really about the ecosystem, the overall support through software to enable the best experience.

A - Atif Malik {BIO 7312618 <GO>}

Great. Colette, if I look at the gaming demand so far this year, a little bit above seasonal due to work or game from home effects.

You guided to a 25% sequential growth for gaming in the October quarter. How do you feel about the holiday season with the new games and broad ray tracing offering?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So when you look at our overall guidance that we provided for our Q3, we are providing guidance that we believe sequentially gaming can grow about 25% from this last quarter, Q2.

We've been doing quite well over the last several quarters, both influenced by just the entertainment area and focus on gaming as COVID-19 has moved people into their home environment. Home environment for them to not only work, for them to learn, but also to play.

Where gaming is continuing to rise and likely the new normal of a very important form of entertainment from the overall home.

Now with the overall launch that we have with our Ampere cards, and we are getting ready for a very important holiday. The holiday season is important because it's also a season where we'll see others enter the market with the overall consoles and a console season for the holiday.

We're expecting a big holiday. A total growth in terms of just the evolution of ray tracing out there with new games that will probably reach this market at this time.

So we're excited for what the second half has for us, not just influenced by our new architecture, but just the season of new games coming to the market for the holiday.

A - Atif Malik {BIO 7312618 <GO>}

Great. Colette, on the console point, how has your PC gaming business fared during prior new console cycles?

We saw Nintendo raising its forecast with the supply chain by 20%.

Obviously you're exposed there. And how committed are you to the console business with Nintendo?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So let me first start with how the market has evolved over generations with the overall consoles. The ecosystem has also changed there. A key part of the driver of the gaming market is just the games. Gamers want to play and overall watch games.

But what is key now is the games are now written to be played across as many platforms as possible.

So the new console ramps that we are expecting raise the bar for overall production of new games. What is very key is that value of games is being inserted with overall ray tracing and getting ready for the season in front of us. Most games right now, therefore, are cross-platform, including many of the most popular franchise. These new console ramps historically have been actually very good for the overall PC gaming universe as well.

RTX support by the new console means it will also have this feature, but great ray tracing content and RTX on NVIDIA is important. The RTX 30 series is a generational leap from where we were just two years ago.

We're the only architecture in market today supporting ray tracing, and we now have two years of improving that overall piece.

So we're excited in terms of what that brings together.

You also asked in terms of our NVIDIA work with Nintendo and Nintendo Switch.

We couldn't be more pleased with that partnership, that partnership of a win-win between both companies and what we have put into the market. They have probably one of the most successful overall consoles in U.S. overall history.

We're committed to continue to help them as we move forward.

A - Atif Malik {BIO 7312618 <GO>}

Great. Colette, how does NVIDIA manage the smooth handover to Ampere from Turing especially given the huge performance to price advantage?

Any color, any comments you can provide on the initial orders for RTX 30 series cards?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So with the launching of our new gaming architecture, it's always a real fun time.

As you've actually noticed with each architecture that we have launched, we've done something differently. The reviews and the commentary on our Ampere architecture

gaming cards have been extraordinary, likely the best launch ever.

We do not have any online orders or preorders at this time.

So we'll take this opportunity to update everybody at the end of the quarter.

A - Atif Malik {BIO 7312618 <GO>}

Great. And just moving on to the gross margin side of the business, you mentioned feeling good about gaining gross margins and growing them generation to generation. While the use of Samsung 8-nanometer was pretty much in line with our expectations for Ampere gaming, with TSMC more on the Ampere data center side, some investors expected TSMC for both. Is sourcing from dual fabs more of a function of supply tightness, pricing or geopolitical risk or everything?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So our decision to use more than one fab provider probably occurred many years ago.

We worked to understand the teams and the processes of both of those fabs, and two of the best fabs, to understand their overall operations. TSMC and Samsung have both been great partners in our success today.

Our decision on which fab to use takes into account a complex set of factors, mostly driven on what is best for the overall architecture or what is best for the overall market platform we are serving, whether or not that is gaming, whether or not it's pro viz or whether or not it's overall data center. Samsung's 8 node has been a special node driven just for overall NVIDIA and really the best overall node that we felt taking to market with Samsung on our overall gaming architecture.

So we're pleased with both of our providers.

But for this one, we are using Samsung to bring it.

A - Atif Malik {BIO 7312618 <GO>}

Great.

You have gained share in the gaming market per the latest Mercury report. NVIDIA is a generation ahead in ray tracing.

One common question I get asked is about growing competition with Big Navi from AMD and then Intel next year. Any kind of thoughts here?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So let's first start with how we had focused on differentiating ourselves in the market, starting even with our prior architectures.

We focused on the most innovative platform, but also matched with unparalleled overall software for the ecosystems and the gamer experience, and that has absolutely served us well.

So really, we are different in terms of how we bring both architectures and our platform to market.

Prior to RTX 30, our overall Turing GPUs with ray tracing and AI technologies were the best GPUs on the market and generations ahead of any of the overall competition's features and/or performance.

We've been now with overall ray tracing in market for two years. Not only does that mean that our cards have ray tracing, that means that we've been working with the ecosystem to support overall ray tracing in many of the games that have come out.

We have a whole host of games that have both been announced and/or in market. Two of the top games that are in market are supported on our overall platform using overall ray tracing, that being Minecraft and now our overall announcement of overall Fortnite.

So our overall GPUs not only incorporate the use of ray tracing, but it also incorporates the use of AI into those GPUs. GPUs that can do DLSS 2.0. It improves the overall performance of that ray tracing by incorporating overall AI to make it even better.

So this is one of our largest generational leaps in performance.

We're up 2x of our performance versus Turing, the second generation of the RTX core, the third generation of the Tensor core.

We feel very good about our competitive position at this time as well as going forward.

A - Atif Malik {BIO 7312618 <GO>}

Great. Let's switch gears to the data center business, now 45% or the largest segment of your company. July quarter was the first full quarter of Ampere ramp. Mellanox was up huge at 30% sequentially, and you expect compute to be the major driver of growth in the October quarter. Where are you in terms of A100 upgrade cycle within hyperscale vertical?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So to ground everybody in terms of what we guided for our data center after just a great overall Q2 being the First Quarter incorporating Mellanox, we couldn't be more pleased on how that is going.

As we move to Q3, yes, we're expecting our data center to grow in the low to mid-single digits sequentially.

On a year-over-year basis, that's about 150% growth. Historically, our data center business stems from both growth in our overall hyperscales, in our enterprises, in our industries.

So keep in mind, with the overall impact on the global economy, we still expect the ramp of the overall Ampere family to grow as we move into Q3.

So let's talk about where we are within that overall Ampere architecture. And you are correct, it is still at the very beginning of its launch.

We began our launch with A100 for the data center focused on the overall hyperscales.

We focused in terms of the A100, which is actually a full system for them, a full system for their overall modern data centers as they continue to build piece by piece rather than a full complete overall compute. This overall system enabled them to quickly qualify and quickly, as you saw, bring up in terms of cloud instances to support the enterprises as well.

We were able to see overall Google reach in terms of its cloud instances.

You've also now seen Microsoft Azure also up with its overall cloud instances.

But also during the quarter, we announced our work in terms of the server manufacturers around the world, now bringing A100 out to the normal server manufacturers out there.

We're expecting 50 different systems to be available between Q3 and Q4, probably about 20 of those systems available in Q2.

So we're still in the early parts of this upgrade cycle for sure.

We're just in the initial quarters, and this is an overall upgrade that will be with us for several years.

A - Atif Malik {BIO 7312618 <GO>}

Great. Colette, a question I get asked is, the cloud ramps that you're seeing at AWS, GCP, Azure, are they offsetting the inventory digestion commentary that we hear from the CPU makers or the storage guys or the GPU model behaves very differently versus these other components?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

We want to step back here a bit and understand that NVIDIA really is about compute adoption, but compute adoption towards acceleration.

So it is different from the normal of overall compute that we've seen for the decades of the past, and we're focused on acceleration as we move forward.

We're new into that product cycle as we've discussed with the overall A100 ramp, and those are very interested in using A100 as we've seen so many changes in the world of AI and acceleration even since the time that we launched the overall V100, our prior architecture.

We're seeing in terms of a 3,000x increase in the size of the overall models as well as the overall complexity of those models. A100 is quite well positioned in that manner to deal with any type of size model, complexity and really work in terms of some of the most important workloads in AI that have evolved over the last couple of years. Those areas in terms of conversational AI, natural language processing and building some of those recommender engines which are so key for those that are online and needing those overall quick responses back.

So we have great visibility into Q3 with our overall hyperscales. What we've seen is we've seen overall mixed results in terms of the vertical industries just given the overall uncertainty in the market.

On-premises can be challenging because of the overall COVID, but we are seeing the industries move towards the cloud.

So we expect a strong Q3 and really focusing on that acceleration, that AI form of compute more than just overall general purpose compute.

A - Atif Malik {BIO 7312618 <GO>}

And when do you see the demand stabilizing on the enterprise on-premise vertical, Colette?

A - Colette Kress {BIO 18297352 <GO>}

So what we have right now is, given the ramp of overall AI00 and as we discussed, starting first with the overall hyperscales and creating those cloud instances that many of our enterprises work, that's a good bridge to begin a quick adoption of the overall enterprises as well as we move forward. Right now and it changes quarter-to-quarter, what we're seeing is the hyperscales are just over 50% of our overall data center revenue.

What we're seeing is the vertical industries are just over 40%, and supercomputing can be less than overall 5%.

So those are important to understand what the mix is and the importance of each of these.

Some of our vertical industries are doing quite well and growing quite well and some of them will probably take some time for them to move to the new architecture given a lot of the challenges that we have with the overall economy.

A - Atif Malik {BIO 7312618 <GO>}

Okay.

On the Mellanox side, Mellanox exceeding expectations so far, what keeps you excited?

A - Colette Kress {BIO 18297352 <GO>}

So we're thrilled that Mellanox is now with us here at NVIDIA. The Mellanox team, I would look at as a perfect match within NVIDIA to join our strategy and our focus in terms of the future of the data center.

Our years of partnership, I think is really key in terms of how that allowed our First Quarter at NVIDIA to work so smoothly.

We are now in the process of beginning our work on the future of the modern data center rather than just getting to know each other.

So we're off to a great start.

We couldn't be more pleased in terms of the work that they have here in front of us.

That modern data center is really about accelerated using the GPU-enabled capabilities, disaggregated using Mellanox-enabled capabilities as well as overall software-defined to really take all the pieces in terms of the data center and accelerate each one of them piece by piece and not just a focus on the time that is overall computing or the GPU is enabling that accelerated computing.

So we're excited to see all the great things working together with Mellanox, but you're correct. It's off to a great start, and we're excited to have them on board.

A - Atif Malik {BIO 7312618 <GO>}

Okay. Colette, this is a tough one.

But again investors kind of always struggle with how to model the growth of the data center business. The team had talked about a \$50 billion TAM by 2023 at its Investor Day last year, split into a 40, 40, 20, hyperscale, enterprise and scientific computing segments. Jensen has talked about cloud being \$100 billion market growing at 40% compounded growth rate. How do you see the long-term data center market growth for NVIDIA?

A - Colette Kress {BIO 18297352 <GO>}

If we step back and look at where we are in the overall ramp of overall AI across the world, AI is still nascent, and it's still advancing exponentially each and every month or essentially each and every year. The model sizes we've talked about, even over three years, have increased 3,000x on the training side before -- and since we've launched overall V100. It has allowed us to open up to new markets as we continue to fuel giant new models that can be trained, focused on natural language processing, probably one of the most difficult pieces of work to understand voice and translate that voice and return back in normal conversation.

Additionally, we have proven the ability to use a programmable GPU, its overall software capabilities to take on a very important market such as inferencing.

Inferencing is a huge market that will take place, not only in the data center, but also on the edge and also in terms of autonomous types of edge devices as we move forward.

So we're still in those early stages of that. With the inclusion of Mellanox, we now have that ability to do more than just accelerating overall compute, but we have an ability to accelerate almost all of the data center with now a focus of the overall networking.

We've likely added tens of billions to our TAM with the inclusion of acceleration and opportunities that we now have with overall Mellanox.

So in the early stages, off to a great start.

But yes, there is a large TAM and a large opportunity for our data center business as a whole going forward.

A - Atif Malik {BIO 7312618 <GO>}

Okay.

This is a tricky one, too. Jensen mentioned on the last earnings call that NVIDIA works closely with ARM, and he respects the energy efficiency of the ARM architecture. Just conceptually, are there advantages or disadvantages, if any, on combining a CPU with a GPU in future compute architectures?

A - Colette Kress {BIO 18297352 <GO>}

So first, I'll start off.

Yes. We work with all CPUs.

We are agnostic to each, any of the overall CPUs that are available. For example, within our DGXs and our new SuperPODs include the AMD CPU as well.

We also work closely with ARM. With each of the workloads, it is a choice of the best overall CPU architecture and is considered when we put some of those workloads together.

Some geographies don't want to be dependent on the x86 overall processors.

So in this case, ARM is the most popular instruction set in the world and essentially fuels billions of devices that are out there. ARM is also very power efficient and is considered for something in the data center to improve that overall data center TCO.

ARM is the only CPU right now without accelerated compute support.

So NVIDIA had started that work in terms of working with ARM to bring CUDA to their instruction set.

We work with a variety of software -- excuse me, a variety of hardware, software companies in the ARM ecosystem.

We believe that CUDA and GPU acceleration for ARM-based servers, such as hyperscales, edge computing and HPC applications, can be with us in the overall future.

So that's how we think about all the different future compute architectures, but our focus is making sure that we are agnostic to all the CPUs that are out there or essentially any of the players that are in the data center in terms of assuring that NVIDIA's overall GPU workloads can work effectively and well to accelerate everything in the data center.

A - Atif Malik {BIO 7312618 <GO>}

Okay. Then switching to auto. Ampere architecture powers next-generation auto platform called Orin, delivering more than 6x the performance of Xavier and 4x

better power efficiency. How does the improved power management profile on Ampere compare with competition?

Will this significantly help to close the gap on Level two and up?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So we're very excited in the last six months what we have brought to market for many of the auto manufacturers, the Tier 1s and many of the overall start-ups.

As you know, we have been focused on an end-to-end platform, enable the cars to take what they have in terms of production of AV from a Level 2, a Level 2+, all the way up to a fully autonomous, and also taking care of what we may want to do in terms of overall robotaxis.

Now an important piece of that is making sure that they have that end-to-end system that enables the ability to advance to Level 3, Level 4, even with the overall same hardware. The energy efficiency is important. But generally, all of our overall GPUs and all of our overall platforms continue to be the best in terms of performance and energy efficiency together in terms of what we're bringing into market.

This last quarter was an important quarter of us to announce one of the most transformational overall deals in the automotive industry.

Our work in terms of with Mercedes is very key and a very important thing that a lot of eyes are on. What it did was take a company, such as Mercedes, and such a rock-solid company that's been available for more than 100 years, and working with us to develop the full software stack that will go into their AV solution. Using our hardware, that the hardware at the onset can continue to upgrade with software and enable many different levels of overall AV over the life of the ownership of the overall car. The great piece of that, which is so great for us as well is the ability to share that revenue with Mellanox, share that software revenue back to overall NVIDIA.

We're not only working inside of the car with our Orin pieces and many other parts of our overall system, but keep in mind that we are working with Mercedes back in their data center to both support leading up to overall production as well as the fleet that will be out there on the road and the data on continuing to work with that data over the time.

So we're really thoughtful in terms of thinking through a supermodel that can be leveraged across the overall industry, but our work with Mercedes is definitely transformational.

A - Atif Malik {BIO 7312618 <GO>}

Great. Then moving to the financial model. Gross margins in second half because of the gaming mix are coming down. Do you see it trending back up to 66% as data center mix normalizes, let's say in first half of next year?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So in our Q2 results, our Q2 results last quarter, we hit record levels at 66% overall non-GAAP gross margin. Remember, the mix of our products still remains our largest driver for our overall gross margins.

Last quarter, with the inclusion of Mellanox, our data center revenue reached record levels and helped drive our overall gross margin expansion. With our Q3 guidance for non-GAAP gross margins, they're just slightly below that record level in Q2. New products that we are introducing have slightly lower margins earlier in the ramp, but we expect those margins to tend to increase as both yields mature and as we optimize a lot of those overall component costs.

So when you think about that overall Ampere architecture as a whole, keep in mind, we have a very strong mix expected of gaming as we move into Q3 from Q2. Within that gaming mix, we also expect all of our platforms to grow, including our consoles.

Our consoles do not have as high a gross margin as the company average.

So in total, we expect to continue to focus on our overall gross margins and focus in terms of launching this Ampere architecture across both data center and gaming.

We expect over the long term to have an ability to move them up over time.

A - Atif Malik {BIO 7312618 <GO>}

Okay.

Then on the OpEx side, hiring and retaining an AI developer, software talent in Silicon Valley is a tough job because you're competing against Internet and software companies. How do you incentivize your employees?

And how should we think about the OpEx growth relative to the sales next year?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So first, talking about our employees and our engineering workforce, both from a software side and a hardware side.

We do have more overall software developers than we actually have hardware developers. The great thing is our software development is done in many places around the world and now is actually done at home as we still have almost all of our sites closed as we are all working from home.

There's no place on the planet that you can go and work like an NVIDIA place. NVIDIA has the ability to produce software for so many different overall workloads and some of the most important workloads of the future. When we think about overall ray tracing for gaming, when you think about everything from virtual reality or augmented reality, enterprise types of graphics, even before we think about the workloads that we are doing in terms of in the data center and the focus on AI and autonomous vehicles. There are some great pieces that we can entice a lot of that talent, but I think what is more important is the talent enjoys NVIDIA.

We have probably one of the best retention that we do in terms of our talent here. People come to work in NVIDIA and do their life's work, their life's work in terms of focus on that GPU overall architecture and thinking about the use of acceleration in so many of these different platforms going forward.

So right now, we feel very good in terms of both our ability to retain and attract talent to overall NVIDIA.

As we turn the corner into next year, as we've just talked about, we've got many very important workloads, very big markets in front of us that will continue to need our focus and continue to see our investment.

So we don't necessarily look at OpEx as a percentage of sales.

We look at it that says, where is the best use of overall OpEx to assure that we can meet these large opportunities in front of us.

So we'll provide more once we turn the corner at the end of this year to focus on what we see as OpEx going forward into fiscal year '22.

A - Atif Malik {BIO 7312618 <GO>}

Okay. Colette, TSMC is building a fab in U.S. next year.

We hear U.S. semiconductor industry lobbying for \$37 billion, with \$17 billion in research and a chip factory. There are geopolitical tensions with China. What does it all mean for NVIDIA?

A - Colette Kress {BIO 18297352 <GO>}

So the idea of bringing a high-end fab here into the U.S., we couldn't be more excited about it.

We work with two of the best overall fab providers on the planet today. Having one of those closer into the U.S. would be great for us.

We deal right now with a worldwide workforce working with those fabs, but having something closer to home here in Santa Clara would also be great for us as well to think about it.

So we couldn't be more supportive of something in that space.

We are working also with the semiconductor industry because I think we need to work together as a group.

We are just one piece of this, and we do support the needs of that overall semiconductor industry.

A - Atif Malik {BIO 7312618 <GO>}

Okay.

We have about three to four minutes left. I'm going to switch to audience questions.

Colette, could you comment further on opportunities for data center for new verticals?

And in particular, for enterprises, could you make some examples?

What is missing for penetration to increase further adoption in enterprises?

A - Colette Kress {BIO 18297352 <GO>}

Yes.

So we've had a great rollout in the last two years focused on building out many of the workloads for the enterprises. What has enabled us to be quite successful in this piece is the focus in terms of the software that will need to be available and the overall production of development resources to enable these workloads as we move forward.

Now we have focused on very key areas.

One, we focused on manufacturing.

We focused on metropolises and building out the overall smart cities.

You've seen us focus in terms of on Clara.

So when you think about our data center platforms, it's not just focused on that hardware.

We have built so much of the software libraries, components, compilers and working with many of the different frameworks that enable it.

We have enabled in terms with our Tensor cores for depending on the workload and depending on that precision in order to address the overall workloads that are out there.

Our focus in terms of e-tail and retailers have been very key as well.

You've seen us focus with large lighthouse retail, such as Walmart, focused on very large distribution, such as the U.S. Postal Service.

So each and every one of these workloads continues to drive overall acceleration into the enterprises.

We also know the enterprises are focused on cloud computing as well.

So we hit in both areas, both in terms of edge computing that we can provide to them, but also what we can provide in those cloud instances.

Let's step back and look at what A100 provides. A100 provides a solution for you to not have to choose or determine at the onset whether or not you need to do training or whether or not you need a virtual VM, okay?

You can do both.

You can do inferencing and training on the same platform. This is great not only for the hyperscales as they built out their data centers, but it is also a very useful technique for the enterprises as they think about the reuse and redistributing that overall compute for the latest types of workloads that we're doing.

So our work is essentially workload by workload or industry-specific and really stems with the application and the software.

One more piece that is important is our work in terms of the data scientists that are out there or the data analytics. This is a very important piece of almost every enterprise, and it's not necessarily industry-specific. The enterprises are out there spending time working with their data scientists to understand how to leverage their data to make overall decisions. It's a relatively inefficient process. They spend a lot of time loading data. They spend a lot of time reanalyzing data, rerunning overall queries to get the information that they were looking for.

We have now been able to stitch ourselves together with Spark, Spark 3.0, which is a very key component in terms of the work that they do. Accelerating that entire process, from end to end in terms of the data load, the working with the data and feeding back out that data, can improve something that is integral to all of the overall enterprises that are out there.

So more to come, but this is what all of our software engineers continue to do every day is more and more workloads that address the industries and the enterprises around the world.

A - Atif Malik {BIO 7312618 <GO>}

Great.

We're almost at the end of time.

Colette, thank you for being part of the Citi conference. Really appreciate it.

A - Colette Kress {BIO 18297352 <GO>}

Okay. Thank you, so much for hosting us.

We'll see you next year as well.

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.