FINAL

Bloomberg Transcript

Q4 2020 Earnings Call

Company Participants

- Colette Kress, Executive Vice President and Chief Financial Officer
- Jensen Huang, President and Chief Executive Officer
- Simona Jankowski, Vice President of Investor Relations

Other Participants

- Aaron Rakers, Analyst
- Atif Malik
- Blayne Curtis, Analyst
- CJ Muse, Analyst
- Harlan Sur, Analyst
- Joe Moore, Analyst
- Mark Lipacis, Analyst
- Matt Ramsay, Analyst
- Timothy Arcuri, Analyst
- Toshiya Hari, Analyst
- Vivek Arya, Analyst
- William Stein, Analyst

Presentation

Operator

And I am your conference operator today. Welcome to NVIDIA's Financial Results Conference Call. All lines have been placed on mute. After the speakers' remarks there will be a question-and-answer period. (Operator Instructions) Thank you. I'll now turn the call over to Simona Jankowski, Vice President of Investor Relations, to begin your conference.

Simona Jankowski {BIO 7131672 <GO>}

Thank you. Good afternoon, everyone and welcome to NVIDIA's conference call for the fourth quarter of fiscal 2020. With me on the call today from NVIDIA are Jensen Huang, President and Chief Executive Officer; and Colette Kress, Executive Vice President and Chief Financial Officer. I'd like to remind you that our call is being webcast live on NVIDIA's Investor Relations website. The webcast will be available for replay until the conference call to discuss our financial results for the first quarter of fiscal 2021. The content of today's call is NVIDIA's property. It can't be reproduced or transcribed without our prior written consent.

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 the disclosure in today's earnings release, 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, February 13, 2020, based on information currently available to us. Except as required by law, we assume no obligation to update any such statements. During this call, we will discuss non-GAAP financial measures. You can find a reconciliation of these non-GAAP financial measures to GAAP financial measures in our CFO commentary, which is posted on our website.

With that, let me turn the call over to Colette.

Colette Kress {BIO 18297352 <GO>}

Thanks, Simona. Q4 revenue was $3.11 billion, up 41% year-on-year and up 3% sequentially, well above our outlook reflecting upside in our data center and gaming businesses. Full year revenue was $10.9 billion, down 7%, we recovered from the excess channel inventory in gaming and an earlier pause in hyperscale spending, and exited the year with great momentum. Starting with gaming revenue of $1.49 billion was up 56% year-on-year and down 10% sequentially. Full year gaming revenue was $5.52 billion, down 12% from a prior year. We enjoyed strong end-demand for our desktop and notebook GPUs.

Let me give you some more details. Our gaming lineup was exceptionally well positioned for the holidays with unique ray tracing capabilities of our RTX GPUs, an incredible performance at every price point. From the Singles' Day shopping event in China through the Christmas season in the West, channel demand was strong for our entire stack. Fueling this were new blockbuster games like Call of Duty: Modern Warfare, continued eSports momentum, and new RTX Super products with RTX price points as low as $299, ray tracing is now the sweet spot for PC gamers. Gaming is thriving, and gamers prefer GeForce. The global phenomenon of eSports keeps gaining momentum with an audience now exceeding $440 million, up over 30% in just two years, according to Newzoo. The League of Legends World Championship brought more than a 100 million viewers on par with this month's Super Bowl. Ray tracing titles continue to come to market and GeForce RTX GPUs are the only ones that support this important technology. This quarter Wolfenstein: Youngblood and Deliver Us The Moon were the latest titles to support ray tracing as well as NVIDIA's deep learning, super sampling technique, which also uses AI to boost performance. With the proliferation of RTX-enabled games and our best ever top to bottom performance, we are solidly into the Turing architecture upgrade cycle. Gamers continue to move to higher-end GPUs to gain better performance and support for ray tracing.

Gaming laptops posted double-digit year-on-year growth for the eighth consecutive quarter. The category continues to expand, driving by appealing, thin and light form factors with fantastic graphics performance. This holiday season retailers stocked a record 125 gaming laptops based on NVIDIA GPUs, up from 94 last year with our Max-Q designs

up two times. At CES, we launched the world's first 14-inch GeForce RTX laptop with Asus. We also continue to expand our Studio line-up of laptops for the fast growing population, our freelance creators, designers and YouTubers, with 13 new RTX Studio systems introduced at CES. Powered by Turing GPUs, these systems are optimized for over 55 creative and design applications with RTX accelerated ray tracing and or AI.

Last week, we launched our GeForce NOW cloud gaming service powered by GeForce. GeForce NOW is the first cloud gaming service to deliver ray-traced games. It's also the only open platform so gamers can enjoy the games they already have and use their existing store accounts without having to repurchase games. GeForce NOW enables PC games on Macs, Windows, PCs, TVs, mobile devices and soon Chromebooks. GFN has a freemium business model that includes 2 membership plans. A free membership with standard access and a founders tier with a starting price of $4.99 per month, which gives priority access and RTX ray tracing support. Our goal with GeForce NOW is to expand GeForce gaming to more gamers. About 80% of GeForce NOW gamers are playing on underpowered PCs or devices with Mac OS or Android. With GeForce NOW, they are able to enjoy PC gaming on a GeForce GPU in the cloud.

GeForce NOW can expand GeForce well beyond the roughly 200 million gamers we reach today. Separately, we entered into a collaboration with Tencent, the world's largest gaming platform to bring PC gaming in the cloud to China, the world's largest gaming market. NVIDIA GPU Technology will power Tencent's START cloud gaming service, which is in early testing stages. Moving to data center. Revenue was a record $968 million, up 43% year-on-year and up 33% sequentially. Our strongest ever sequential growth in dollar terms. Full-year fiscal year '20 datacenter revenue was a record $2.98 billion, up 2% from a prior year. Strong growth was fueled by hyperscale, and vertical industry and customers. Hyperscale demand was driven by purchases of both our training and inference products in support of key AI workloads such as natural language understanding, conversational AI and deep recommendaters [ph]. Hyperscale demand was also driven by cloud computing, AWS now makes the T4 available in every region.

This underscores the versatility of the T4 which excels at a wide array of high-performance computing workloads, including AI inference, cloud gaming, rendering and virtual desktop. Vertical industry growth was driven primarily by consumer Internet companies. Other verticals such as retail, health care and logistics, continue to grow from early stage build outs with a strong foundation of deep learning engagements, and we see an expanding set of opportunities across high performance computing, data science and edge computing applications. T4 our inference platform had another strong quarter with shipments up four times year-on-year driven by public cloud deployments, as well as edge AI video analytics applications.

T4 and V100 reflecting strong demand for inference and training respectfully set records this quarter for both shipments and revenue. Even as NVIDIA remains the leading platform for AI model training, NVIDIA's inference platform is getting wide use by some of the world's leading enterprise and consumer Internet companies including American Express, Microsoft, PayPal, Pinterest, Snap and Twitter.

FINAL

Bloomberg Transcript

The industry continues to do groundbreaking AI work for NVIDIA, for example Microsoft's biggest quality improvements made over the past year in its Bing search engine stem from its use of NVIDIA GPUs and software for training and inference of its natural language understanding models. These DNN Transformer models popularized by BERT have computational requirements for training that are in the order of magnitude higher than earlier image-based models. Conversational AI is a major new workload requiring GPUs for inference to achieve high throughput within the desired low-latency. Indeed, Microsoft cited an inference throughput increase of up to 800 times on NVIDIA GPUs compared with CPUs, enabling it to serve over one million BERT influences per second worldwide.

And just this week, Microsoft researchers announced a new breakthrough in natural language processing with the largest ever publicized model trained on NVIDIA DGX-2. This advances the state-of-the-art for AI assistance in task such as answering questions, summarization and natural language generation. Recommendaters are also an important machine learning model for the Internet, powering billions of queries per second. The industry is moving to deep recommendaters such as wide and deep model which leverage deep learning to enable automatic feature learning and to support unstructured content. Running these models on GPUs can dramatically increase inference throughput and reduce latency compared with CPUs, for example, Alibaba's and Baidu's recommendation engines run on NVIDIA AI boosting their inference throughput by orders of magnitudes beyond CPUs. Deep recommendaters enabled Alibaba to achieve 10% increase in click-through rates. We also announced the availability of a new GPU accelerated supercomputer on Microsoft Azure. It enables customers for the first time to rent an entire AI supercomputer on demand from their desk, matching the capabilities of large on-premise supercomputers that can take months to deploy. And in Europe, energy company ENI announced the world's fastest industrial supercomputer based on NVIDIA GPUs.

AI has even come to pizza delivery. At the National Retail Federation's Annual Conference last month we announced Domino's as a customer deploying our platform for deep learning and data science applications, helping with customer engagement and order accuracy prediction. More broadly in retail we have seen a significant increase in the adoption of NVIDIA's edge computing offerings by large retailers for powering AI applications that reduce shrinkage, optimize logistics and create operational efficiencies.

At the SC19 Supercomputing Conference, we introduced a reference design platform for GPU accelerated ARM-based servers along with ecosystem partners ARM and peer computing Fujitsu and Marvell. We made available our ARM compatible software development kit consisting of NVIDIA CUDA-X libraries and development tools for accelerating computing. This opens the floodgates of innovation to support growing new applications from hyperscale cloud to access sales supercomputing.

We also introduced NVIDIA Magnum IO, a suite of software optimized to eliminate storage and input-output bottlenecks. Magnum IO delivers up to 20 times faster data processing for multi-server, multi-GPU computing nodes when working with massive data sets to carry out complex financial analysis, climate modeling and other workloads for data scientists, high-performance computing and AI researchers. Finally, we introduced TensorRT7, a

seventh generation of our inference software development kit, which speeds up components of conversational AI by 10 times comparing to running on CPUs. This helps drive latency below the 300 milliseconds threshold considered necessary for real-time interactions supporting our growth in conversational AI. Moving to ProViz. Revenue reached a record $331 million, up 13% year-on-year and up 2% sequentially.

Full year revenue was a record $1.21 billion, an increase of 7% from the prior-year. ProViz accelerated in Q4 as the rollout of more RTX-enabled applications is driving strong upgrade cycle for our Turing GPUs. RTX is also opening up new market segment opportunities such as rendering and Studio for freelance creators. In November, Vray, Arnold and Blenders software vendors began shipping with RTX technology. These join our leading creative and design applications including Premier Pro, Dimension, SOLIDWORKS, Catia and Maya.

With RTX, these applications enable enhanced creativity and notable productivity gains. In Blenders Cycles, for example, real-time rendering performance is boosted four times versus a CPU. RTX is now supported by more than 40 leading creative and design applications, reaching a combined user base of over 40 million.

Finally, turning to automotive. Revenue was $163 million, flat from a year ago and up 1% sequentially. Full year revenue reached a record $700 million, up 9% year-on-year. During the quarter, we announced DRIVE AGX Orin, the next generation platform for autonomous vehicles and robots, powered by our new Orin SoC and delivering nearly seven times the performance of the previous generation Xavier SoC. The platform scales from Level 2 plus AI Assisted driving up to Level 5 fully driverless operation. Orin is software defined and compatible with Xavier allowing developers to leverage their investment across multiple product generations.

Moving to the rest of the P&L. Q4, our GAAP gross margins were 64.9% and non-GAAP was 65.4%, up sequentially, largely reflecting a higher contribution of data center products. Q4 GAAP operating expenses were $1.02 billion and non-GAAP operating expenses were $810 million, up 12% and 7% year-on-year, respectively. Q4 GAAP EPS was $1.53, up 66% from a year earlier. Non-GAAP EPS was a $1.89, up 136% from a year ago.

Q4 cash from operations was $1.46 billion. Fiscal year '20, cash flow from operations was a record $4.76 billion.

With that, let me turn the outlook for the first quarter of fiscal 2021. The outlook does not include any contribution from the pending acquisition of Mellanox. We are engaged and progressing with China on the regulatory approval and believe the acquisition will likely close in the first part of calendar 2020. Before we get to the new -- the numbers, let me comment on the impact of the Coronavirus. While it is still early and the ultimate effect is difficult to estimate, we have reduced our Q1 revenue outlook by $100 million to account for the potential impact. We expect revenue to be $3 billion, plus or minus 2%. GAAP and non-GAAP gross margins are expected to be 65% and 65.4%, respectively, plus or minus 50 basis points.

GAAP and non-GAAP operating expenses are expected to be approximately $1.05 billion and $835 million, respectively. GAAP and non-GAAP OI, in year [ph], both is expected to be income of approximately $25 million. GAAP and non-GAAP tax rates are both expected to be 9%, plus or minus 1%, excluding discrete items. Capital expenditures are expected to be approximately $150 million to $170 million.

Further financial details are included in the CFO commentary and other information available on the IR website. In closing, let me highlight an upcoming event for the financial community. We will be at the Morgan Stanley Technology, Media and Telecom Conference on March 2 in San Francisco.

With that, we will now open the call for questions. Operator, will you please poll for questions?

Questions And Answers

Operator

(Operator Instructions) And our first question comes from the line of Toshiya Hari with Goldman Sachs.

Q - Toshiya Hari {BIO 6770302 <GO>}

Hi guys. Thanks very much for the question, I guess, on data center, Colette Or Jensen, can you speak to some of the areas that drove the upside in the quarter. You talked about inference, and both the T4 in the V100 having record quarters but relative to your internal expectations, what were some of the businesses that drove the upside. And if you can also speak to the breadth of your customer profile today relative to a couple of years ago, how that's expanding, that would be helpful as well? Thank you.

A - Jensen Huang {BIO 1782546 <GO>}

Yeah Toshiya. Thanks a lot for your question. The primary driver for growth is AI. There are four fundamental dynamics, the first is that the -- the AI models that are being created are achieving breakthroughs, and quite amazing breakthroughs, in fact. In natural language understanding, in conversational AI, in recommendation systems, and you know this -- you know this, but for the others -- others in the audience, recommendation systems are essentially the engine of the Internet today. And the reason for that is because there are so many items in the world. Whether it's a store or whether it's content or websites or information you're querying, there are -- there are hundreds of billions, trillions and depending how you count it, hundreds of trillions of items in the world.

And there are billions of people, each with their own characteristics, and their countless contexts, and between the items, the people, the users and the various contexts that we're in, location and what you're looking for, and weather or what's happening in the environment. Those kind of context affects the search query, the answer they provide you. The recommendation system is just foundational now to search. And some people have said this the end of search, and the beginning in the era of recommendation systems.

Work is being done everywhere around the world in advancing recommendation systems. And very first time, over the last year, it's been able to be done in deep learning. And so the first thing is just the breakthroughs in AI. The second is production AI, which means that whereas we had -- we had significant, and we continue to have significant opportunities in training, because of the models getting larger there are more of them, we're seeing a lot of these models going into production, and that business is called inference. Inference as Colette mentioned grew four times year-over-year as a substantial part of our business now. But one of the interesting statistics is TensorRT7, the entire TensorRT download this year was about 500,000, a doubling over a year ago. Where most people don't understand about inference is -- is an incredibly complex comp -- computational problem, but it's an enormously complex software problem.

And so the second -- the second dynamic is moving from training or growing from training and models going into production called inference. The third is the growth, not just in hyperscale anymore, but in public -- in public cloud and in vertical industries public cloud because of thousands of AI start-ups that are now developing AI software in the cloud, and the OpEx model works much better for them as they're younger. When they become larger, they could decide to build their own data center infrastructure on-prem, but the thousands of startups start their lives in the cloud.

We're also seeing really great success in verticals. One of the -- one of the most most exciting vertical is logistics, logistics, retail, warehousing. We announced, I think this quarter or last -- end of last quarter, USPS, American Express, Walmart, just large companies who have enormous amounts of data that they're trying to do data analytics on, and to do predictive analytics on. And so that the third dynamic is the growth in beyond hyperscale in public cloud as well as vertical industries. And then, the last dynamic is being talked about a lot and this is really, really exciting, and it's called edge AI. We used to call it industries and AI where the action is, but the industry, it now calls it edge AI. We're seeing a lot of excitement there. And the reason for that is, is that you need to have low latency inference, you might not be able to stream the data all the way to the cloud for cost reasons or data sovereignty reasons, and -- and you need the response time. And so those four dynamics around AI really drove our growth.

Operator

Your next question comes the line of Joe Moore with Morgan Stanley.

Q - Joe Moore {BIO 17644779 <GO>}

Great, thank you. Just following up on that, as you look back at the last 12 months and the deceleration that you saw in your HPC cloud business, now that you have the perspective of seeing what's driving the rebound. Any thoughts on what drove it to slow down in the first place, was it just digestion. Was it sort of a hand off from image recognition to these newer applications that you just talked about. Just help us what happened there and I guess, as it -- as it pertains to the future, do we think of this as a business that will have that kind of lumpiness to it?

A - Jensen Huang {BIO 1782546 <GO>}

Yeah, that's a really good question. In fact, if you look backwards, now we have the benefit of history. The Deep recommendation systems, the natural language understanding breakthroughs, the conversational AI breakthroughs all happened in this last year. And the velocity by which the industry captured the benefits here and continue to evolve and advance from these what, so-called transformer models was really quite incredible.

And so the -- all of a sudden, the number of breakthroughs in AI has just grown tremendously and these models have grown tremendously. Just this last week. Microsoft announced that they've trained -- their new-and-knit model in collaboration with the work that we did, we call Megatron, increased the size of a model from 7.5 billion parameters to 17.5 billion parameters.

And the the accuracy of their natural language understanding has just -- has really been boosted. And so models are -- are -- AI's finding really, really fantastic breakthroughs and models are getting bigger, and there are more of them. And when you look back and look at when these these breakthroughs happened, it essentially happened this last year. The second, we've been working on inference for some time, and until this last year, very few of those inference models went into production and now we have deep learning models across all of the hyperscalers in production. And this last year, we saw really great growth in inference.

The 3rd dynamic is probably clouds. All these AI start-ups that are being started all over the world's about 6,000 of them. They're starting to -- they're starting to develop and be able to put their models into production. And with the scale out of AWS, we now have T4s in every single geography. So the combination of of the availability of our GPUs in the cloud and these startups and vertical industries deploying their AI models into production. The combination of all that, just kind of came together and all of that happened this last year, and as a result, we had record sales of V100s and T4s. And so, we're quite excited with the developments and it's all really powered by AI.

Operator

Your next question comes from the line of Vivek Arya with Bank of America Securities.

Q - Vivek Arya {BIO 6781604 <GO>}

Thanks for taking my question and congratulations on returning the business back to the strong growth. Jensen, I wanted to ask about how you are positioned from a supply perspective for this coming yea. Your main foundry is running pretty tight. How will you be able to support the 20% or so growth year that many investors are looking for. If you could just give us some commentary on how you're positioned from a supply perspective, that'll be very helpful.

A - Jensen Huang {BIO 1782546 <GO>}

Well, I think we're in pretty good shape on supply. We surely won't have ample supply, it is true that the industry is tight and the combination of supporting multiple processes, multiple fabs across -- across our partner TSMC, we've got a lot of different factories and a lot of different -- several different nodes of process qualified. I think we're in good shape.

And so, so we just have to watch it closely, and we're working very closely with all of our customers and forecasting, and of course, that gives us better visibility as well. And -- but all of us have to do a better job forecasting and we're working very closely between our customers and our foundry partners, TSMC.

Operator

Your next question comes from line of Timothy Arcuri with UBS.

Q - Timothy Arcuri {BIO 3824613 <GO>}

Hi, thanks. Colette, I'm wondering if you can give us, in data center, if you can give us a little idea of what the mix was between industries and hyperscale. I think last quarter hyperscale was a little bit less than 50%. Can you give us maybe the mix or how much it was up, something like that? Thanks.

A - Colette Kress {BIO 18297352 <GO>}

Yeah. Tim, thanks for the question. Similar to what we had seen last quarter, with all things growing as we moved into this quarter, growth in terms of the hyperscales continued expansion in terms of those vertical industries, and even in the cloud instances. We're still looking at around the same split of 50-50 between our hyperscales and our vertical industries and maybe a little bit tad below 50 in terms of our total overall hyperscales.

Operator

Your next question comes from the line of Aaron Rakers with Wells Fargo.

Q - Aaron Rakers {BIO 6649630 <GO>}

Yeah, thanks for taking the question, and congratulations on the results. When I look at the numbers, the growth on an absolute basis sequentially in data center was almost two times or north of two times -- what we've seen in the past as far as the absolute sequential change. Through the course of this quarter, you were pretty clear that you would expect to see an acceleration of growth in the December quarter. I'm just curious of how you think about that going into the April quarter, and how we should think about that growth rate through the course of this year. If you can give us any kind of framework. And Jensen, just curious, I mean, as you think about the bigger picture, where do you think we stand from an industry perspective today in terms of the amount or the attach rate of GPUs, is it for acceleration in the server market. Where do you think that might be looking out over the next three years or so. Thank you.

A - Jensen Huang {BIO 1782546 <GO>}

Thanks, Aaron. Colette, do you want to go first?

A - Colette Kress {BIO 18297352 <GO>}

Sure. When we think about going into Q1 and our data center overall growth, we do expect to see continued growth, both going into Q1. We believe our visibility still remains positive, quite well and we're expecting that as we move into it and go forward.

A - Jensen Huang {BIO 1782546 <GO>}

Yeah, Aaron. I believe that every query on the Internet will be accelerated someday. And at the very core of it most, almost all queries will have some natural language understanding component to it, almost all queries will have to sort through and make a recommendation from the trillions of possibilities, filter it down and recommend a handful of recommended answers to your queries, whether it's -- whether it's shopping or movies or just asking locations or even asking a question. The number or the possibilities of all the answers versus what is best answer is -- needs to be filtered down. And that filtering process is called recommendation. That recommendation system is really complex. And deep learning is going to be involved in all that. That's the first thing. I believe that every query will be accelerated. The second is -- is as you know, CPU scaling has really slowed and there's just no two ways about it. It's not a marketing thing, it's a physics thing.

And the ability for CPUs to continue to scale without increasing costs or increasing power has ended. And it's called the end of Dennard scale. And so, there has to be another approach. The combination of -- of the emergence of deep learning and the use of artificial intelligence, and the amount of computation that's necessary to -- for every single query but the benefit that comes along with that, and the end of Dennard scaling suggested there needs to be another approach, and we believe that approach is acceleration.

Now our approach for acceleration is fundamentally different than the -- than an accelerator. Notice, we never say accelerator, we say accelerated computing and the reason for that is because we believe that a software defined data center will have all kinds of different AIs. The AIs will continue to evolve, the models will continue to evolve and get larger and the software defined data center needs to be programmable.

It is one of the reasons why we've been so successful and if you go back and think about all the questions that have been asked of me over the last three or four years around this area, the consistency of the answer has to do with the programmability of architecture, the richness of the software, the difficulties of the compilers, the ever-growing size of the models, the diversity of the models and the advances that these models are creating.

And so we're seeing the beginning of a new computing era. And a fixed function accelerator is simply not the right answer. And so, we believe that the future is going to be accelerated. It's going to require an accelerated computing platform, and software richness is really vital so that these data centers could be software defined.

And so, I think that we're in the early innings. The early innings, very, very early innings of this new future, and I think that accelerated computing is going to become more and more important.

Operator

Your next question comes from the line of Matt Ramsay with Cowen.

Q - Matt Ramsay {BIO 17978411 <GO>}

Thank you very much. Good afternoon. And obviously, congratulations on the data center success. I wanted to ask a little bit, Colette, about the -- you took $100 million out for Coronavirus, and I wanted to ask a little bit about how you got to that number. Really two pieces, one, if you could remind us, maybe in terms of units or revenue, how, what percentage of your gaming business is within China. And as you looked at that $100 million that you put out of the guidance, are you thinking about that from a demand destruction perspective or are you thinking about it from something in the supply chain that might limit your sales? Thank you.

A - Colette Kress {BIO 18297352 <GO>}

Sure. Thanks for the question, Matt. So it's really still quite early in terms of trying to figure out what the impact from the overall Coronavirus may be. So we're not necessarily precise in terms of our estimates. Yes, our estimates are split between an impact possibly on gaming and data center, and split pretty much equally. The $100 million also reflects what may be supply challenges or maybe overall demand. But we're still looking at those to get a better understanding where though we think that might be. In terms of our business and our business make up. Yes. Our overall China business for gaming is an important piece -- where about 30% of our overall China gaming as a percentage of our overall gaming business.

For data center, it's -- it moves quite a bit. They are very important market for us, but it moves from quarter-to-quarter, just based on the overall end-customer mix as well as the system builders that they may choose. So it's a little harder to determine.

Operator

Your next question comes from the line of Harlan Sur with JPMorgan.

Q - Harlan Sur {BIO 6539622 <GO>}

Good afternoon and congratulations on the strong results and guidance. (Multiple Speakers) On the cloud gaming. Yeah, no problem. Good to see the recent launch of your GeForce NOW service. But on the partnership with Tencent on cloud gaming, it sounds like Tencent should have a smoother transition to the cloud model. They are the largest gaming company in the world. So they own many of the games. They also have their own data center infrastructure already in place, but how is the NVIDIA team going to be supporting this partnership. Is it going to be near [ph] your GeForce NOW hardware framework or will you just be supporting them with your standalone GPU products, and when do you expect the service to go mainstream?

A - Jensen Huang {BIO 1782546 <GO>}

Let's see. Tencent is the world's largest publisher. China represents about a third of the world's gaming, and transition to the cloud is going to be a long-term journey. And the reason for that is because Internet connection is not consistent throughout the entire market, and a lot of application still needs to be on-boarded. And we're working very closely with them, we're super enthusiastic about it. If we're successful long term and

we're talking about -- we're talking about an extra billion, a billion gamers that we might be able to reach. And so I think that this is an exciting opportunity, just a long-term journey.

Now here in the West, we've had a lot more opportunity to refine the connections around the world and working through the data centers, the local hubs as well as people's WiFi routers at home. And so we've been in beta for quite some times as you know, and here in the West our -- our platform is open and we have several hundred games now, and we're in the process of on-boarding another 1500 games. We're the only cloud platform that's based on Windows and allows us to be able to bring PC games to the cloud. And so the reach is, we've had more experience here in the west with reach and we've had -- we obviously have a lot more games that we can on-board.

But I'm super enthusiastic about the partnership we have with Tencent. Overall, our GeForce NOW, you guys saw the launch. It's -- the reception has been fantastic, reviews have been fantastic. Our strategy has three components. So there is the -- the GeForce NOW service that we provide ourselves. We also have GeForce NOW alliances with telcos around the world to reach the regions around the world that we don't have a presence in, and that is going super well and I'm excited about that. And then, lastly partnerships with large publishers for example like Tencent. And we offer them our platform, of course, and a great deal of software and just a lot of engineering that has to -- that has to be done in collaboration to refine the service.

Operator

Your next question comes from the line of CJ Muse with Evercore.

Q - CJ Muse

Yeah, good afternoon. Thank you for taking my question. I guess, a question on the gaming side, if I look at your overall revenue guide it would seem to suggest that you're looking for, typically, I guess, better seasonal trends into April, and I guess, can you speak to that. And then, how are you seeing desktop gaming demand with increasing content becoming more available. How should we think about the growth trajectory through 2020. And then, just really as a modeling question as part of gaming, with notebook now a third of the revenues, how should we think about kind of the seasonality going into April and July for that part of your business? Thank you.

A - Jensen Huang {BIO 1782546 <GO>}

Yeah. So CJ, I'm going to go first and then Collette is going to take it home here. So the first part of it is this, our gaming business has -- the end -- I'm sorry? Okay. Our gaming business, the end-market demand is really terrific. It's really healthy. It's been healthy throughout the whole year. And -- and it's pretty clear that RTX is doing fantastic, and it's very -- it's super clear now that ray tracing is the most important new feature of next-generation graphics. We have 30 -- over 30 games that are -- that have been announced. 11 games or so that have been shipped. The pipeline of ray tracing games that are going to be coming out is just really, really exciting. The second factor -- and one more thing about RTX, we finally have taken RTX down to $299. So, it's now at the sweet spot of gaming.

FINAL

Bloomberg Transcript

And so, RTX is doing fantastic, its sell-through is fantastic all over the world. The second part, the second part of our business that is changing in gaming, is this the amount of notebooks sales and the success of Nintendo Switch has really changed the profile of our overall gaming business. Our notebook business, as Colette mentioned earlier, has seen double-digit growth for eight consecutive quarters, and this is unquestionably a new gaming category, like it's a new game console. This is going to be the largest game console in the world, I believe.

And the reason for that is because there are more people with laptops and there -- of any other device. And so the fact that we've been able to get RTX into a thin and light notebook -- a thin and light notebook is -- is really a breakthrough, and it's one of the reasons why we're seeing such great success in notebook. Between the notebook business and our Nintendo Switch business, the profile of gaming overall has changed and it's become more seasonal. It's more seasonal because -- because devices, systems like notebooks and Switch are built, largely in two quarters, Q2 and Q3.

And they build it largely in Q2 and Q3 because it takes a while to build them and ship them, and put them into the hubs around the world, and they tend to build it ahead of the holiday season. And so that's one of the reasons why Q3 will tend to be larger, and Q4 will tend to be more seasonal. Q1 will tend to be more seasonal than the past. But the end demand is fantastic. RTX is doing great, and part of it is just a result of the success of our notebooks. I'm going to hand it over to Colette.

A - Colette Kress {BIO 18297352 <GO>}

Yeah. So with that, from a background, and you think about all those different components that are within gaming, the notebook, the overall Switch and, of course, all of the ray tracing that we have in terms of desktop. Our normal seasonality, as we look at Q1 for gaming with all those three pieces is usually sequentially down from Q4. Sequentially down Q4 to Q1. This year the outlook assumes it will probably be a little bit more pronounced due to the Coronavirus. So in total, we're probably looking at Q1 to be in the low double-digit sequential decline in gaming.

Operator

Your next question comes from the line of Atif Malik with Citi.

Q - Atif Malik {BIO 7312618 <GO>}

Hi. Thank you for taking my question and good job on results and guide. On the same topic, Coronavirus, Colette, I'm a bit surprised that the guidance, the range on the guidance is not wider versus historic. Can you just talk about why not widen the range and what went into that $100 million hit from the Coronavirus.

A - Colette Kress {BIO 18297352 <GO>}

So Atif, thanks for the question. Again, it's still very early regarding the Coronavirus. Our thoughts are out with both the employees, the families and others that are in China. So our discussions both with our supply chain that is very prominent in the overall Asia region as well as our overall AIC makers, as well as our customers is as about as timely as we can

be, and that went into our discussion and our thoughts on the overall guidance that we gave into a $100 million. We'll just have to see how the quarter comes through, and we will discuss more when we get to it, but at this time, that was our best estimate at this time.

Operator

Your next question comes from the line of William Stein with SunTrust.

Q - William Stein {BIO 15106707 <GO>}

Great, thanks for taking my question. Jensen, I'd love to hear your thoughts as to how you anticipate the inference market playing out. Historically, NVIDIA's had essentially all of the training market and little of the inference market in the last year-and-a-half or so. I think that's changed, where you've done much better in inference. Now you have the T4 in the cloud. You have EGX at the Edge, and you have Jetson, I think is what it's called, at the sort of endpoint device. How do you anticipate that market for inference developing across those various positions and how are you aligning your portfolio for that growth?

A - Jensen Huang {BIO 1782546 <GO>}

Yeah. Thanks a lot Will. Let's see, I think, the -- historically, inference has been a small part of our business because AI was still being developed. Deep learning -- AI is not -- historical AI, classical machine learning weren't particularly suited for GPUs, and weren't particularly suited for acceleration. It wasn't until deep learning came along that the amount of computation necessary, it's just extraordinary and the second factor is the type of AI models that were developed, eventually, it -- the type of models related to natural language understanding and conversational AI and recommendation systems. These required instantaneous response, the faster the answer, the more likely someone is going to click on the answer.

And so, you know that latency matters a great deal and it's measurable. The effect on the business is historically measured. And so, for conversational AI, for example, we've been able to reduce the latency of the entire pipeline from speech recognition to the language processing to -- for example, fix the errors and -- and such, come up with a recommendation to text-to-speech, to the voice synthesis, that entire pipeline could take several seconds. We run it so fast that it's possible now for us to process the entire pipeline within a couple of 100 milliseconds to 300 milliseconds.

That is in the realm of interactive conversation. Beyond that is just simply too slow. And so, the combination of AI models that are large and complex, that are moving to inference, moving to production. And then secondarily, conversational AI and latency sensitive models and applications, where our GPUs are essential, now moving forward. I think you're going to see a lot more opportunities for us in inference. The way to think about that long term is acceleration is essential, because of -- end of Dennard scaling. Process technology is going to reap [ph] demand that we compute in a different way and the way that AI has evolved, the deep learning, it suggests that acceleration on GPUs is just a really phenomenal approach.

Data centers are going to have to be software defined. And I think, as I mentioned, I think I mentioned earlier to another question, I believe that in the future, the data center will all be accelerated, it will all be -- it will be all running AI models and it will be software defined. And will be programmable, and having an accelerated computing platform is essential.

As you move out to the Edge. It really depends on whether your platform is software defined, whether it has to be programmable or whether it's fix functioned. There are many, many, many devices where the inference work is very specific. It could be something as simple as detecting changes in temperature or changes in sound or detecting motion.

Those type of -- those type of inference models are -- could still be based on deep learning. It's function specific. You don't have to change it very often and you're running one or two models at any given point in time. And so those devices are going to be incredibly cost effective. I believe those AI Chips, you're going to have AI chips that are $0.50, a $1, and you're just going to put it into something and -- and it's going to be doing magical detections.

The type of platforms that we're in such as self-driving cars and robotics. The software is so complicated, and there is so much evolution to come yet, and it's going to constantly get better. Those software defined platforms are really the ideal targets for us.

And so we call it AI at the Edge, Edge computing devices. One of the Edge computing devices I'm very excited about is what people call mobile Edge, or basically 5G telco Edge. That -- that data center will be programmable. We recently announced that we partnered with Ericsson and we're going to be accelerating to 5G stack. And so that -- that needs to be a software defined data center, it runs all kinds of applications, including 5G, and and those applications are going to be -- those opportunities are fantastic for us.

Operator

Your next question comes from the line of Mark Lipacis with Jefferies.

Q - Mark Lipacis {BIO 2380059 <GO>}

Thanks for taking my question. Jensen, I guess, I had a question about your -- how you think about the sustainability of your market position in the data center. And I guess in my simplistic view, about 12 years ago you made out a consensus call to invest in CUDA software, distributed to universities. Neural networking took off and you were the de facto standard, and here we are right now. And for me, what's interesting to hear is that the demand that you're seeing today for your products is from markets that just developed within the last year. And my question is like how do you think about your investment -- your R&D investment strategy to make sure that you are staying way ahead of the market, of the competition and even your customers who are investing in these markets, too? Thank you.

A - Jensen Huang {BIO 1782546 <GO>}

Yeah. Thanks, Mark. Our company has to live 10 years ahead of the market. And so, we have to imagine where the world is going to be in 10 years' time, in five years time and work our way backwards. Now, our company is focused on one singular thing, the simplicity of it is incredible, and that one singular thing is accelerated computing -- accelerated computing. And -- and accelerated computing is all about the architecture, of course. It's about the complicated systems that we're in, because throughput is high, when our acceleration we can -- when we can compute 10, 20, 50, a100 times faster than a CPU, all of a sudden everything becomes a bottleneck, memories a bottleneck, networking is a bottleneck, storage is a bottleneck, everything's a bottleneck.

And so we have to be -- NVIDIA has to be a supremely good system designer but the complexity of our of stack, which is the software stack above it, is really where the investments over the course of last some 29 years now has really paid off. NVIDIA's, frankly, has been an accelerated computing company since the day it was born. And -- and so we -- we -- our company is constantly trying to expand the number of applications that we can accelerate. Of course computer graphics was an original one, and we're reinventing it with real time ray tracing. We have rendering which is a brand new application, that we're making great progress in. We just talked -- I just mentioned 5G acceleration. Recently, we announced genomics computing. And so, those are new applications that are really important to the future of computing.

In the area of artificial intelligence, from image recognition to natural language understanding to conversation, to recommendation systems to robotics and animation. The number of applications that we're going to accelerate in the field of AI is really, really broad. And each one of them are making tremendous progress and getting more and more complex.

And so, the question about the sustainability of our company really comes down to two dimensions. Let's assume for the fact, let's assume for now that accelerated computing is the path forward, and we certainly believe so. And there's a lot of evidence from the laws of physics to the laws of computer science that it would suggest that -- that accelerated computing is the right path forward. But there's really, basically, comes down to two dimensions. One dimension is are we continuing to expand, are we continuing to expand the number of applications that we can accelerate.

Whether it's AI or computer graphics or genomics or 5G, for example. And then the number -- and then the second is, those applications, are they getting more impactful and adopted by the ecosystem, the industry. And are they continuing to be more complex. Those dimensions -- the number of applications and the rich -- and the impact of those applications and the evolution. The growth of complexity of those applications, if those dynamics continue to grow, then I think -- I think, we're going to do a good job. We're going to sustain, and so -- and I think -- I think when I spelled it out that way, it's basically the equation of growth of our company. I think it's fairly clear that the opportunities are fairly exciting ahead.

Operator

Your next question comes from the line of Blayne Curtis with Barclays.

Q - Blayne Curtis {BIO 15302785 <GO>}

Thanks for squeezing me in. I just wanted to ask on the auto side. I think at least one of your customers might have float [ph] out their program. Just kind of curious as you look out the next couple of years, the challenges with the OEMs moving slower. And then, just any perspective on the regulatory side, has anything changed there, would be helpful. Thanks.

A - Jensen Huang {BIO 1782546 <GO>}

I think that the automotive industry is struggling, but for all of the reasons that everybody knows. However, the enthusiasm to redefine and reinvent their business model has never been greater. Every single one of them -- every single one of them would know now, and they surely --they've known for some time and autonomous capabilities, is really the vehicle to do that, they need to be tech companies. Every car company wants to be a tech company, they need to be a tech company. Every car company needs to be software defined and the platform by which to do so is an electric vehicle with autonomous autopilot capability. That car has to be software defined and this is their future, and they are racing to get there. And so, although the automotive industry is struggling in near term, their opportunity has never been better, in my opinion. The future of AV is more important than ever. The opportunity is very real, the benefits of autonomous is, for whether it's safety, whether it's utility, where it's cost reduction and productivity, has never been more clear.

And so, I think that I'm as enthusiastic as ever about the autonomous vehicles, and the projects that we're working on are moving ahead. And so the near-term -- near term challenges of the automotive industry or whatever sales slowdown in China that they're experiencing, I feel badly about that. But the industry is as clear-headed about the importance of AV as ever.

Operator

I will now turn the call back over to Jensen for any closing remarks.

A - Jensen Huang {BIO 1782546 <GO>}

We had an excellent quarter with strong demand for NVIDIA RTX graphics and NVIDIA AI platforms, and record data center revenue. NVIDIA RTX is reinventing computer graphics and the market's response is excellent, driving a powerful upgrade cycle in both gaming and professional graphics, while opening whole new opportunities for us to serve the huge community of independent creative workers and social content creators, and new markets in rendering and cloud gaming. Our data center business is enjoying a new wave of growth powered by three key trends in AI, natural language understanding, conversational AI. Deep recommenders, are changing the way people interact with the internet. The public cloud demand for AI is growing rapidly, and as AI shifts from development to production, our inference business is gaining momentum.

We'll be talking a lot more about these key trends and much more at next month's GTC Conference in San Jose. Come join me, you won't be disappointed. Thanks everyone.

Operator

Ladies and gentlemen, this concludes today's conference call. Thank you for participating. You may now disconnect.

FINAL

Bloomberg Transcript