

Q1 2020 Earnings Call

Company Participants

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

Other Participants

- Aaron Rakers
- CJ Muse
- Chris Castle
- Harlan Sur
- Joe Moore
- Matt Ramsay
- Stacy Rasgon
- Timothy Arcuri
- Toshiya Hari
- Vivek Arya

Presentation

Operator

Good afternoon. My name is Christina, and I'll be 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). And I'll now turn the call over to Simona Jankowski from Investor Relations to begin your conference.

Simona Jankowski {BIO 7131672 <GO>}

Thank you. Good afternoon, everyone. And welcome to NVIDIA's Conference Calls for the First 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 second quarter of fiscal 2020. 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, May 16, 2019, 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. Q1 revenue was \$2.2 billion, in line with our outlook and down 31% year-on-year and up 1% sequentially.

Starting with our gaming business. Revenue of \$1.05 billion was down 39% year-on-year and up 11% sequentially, consistent with our expectations. We are pleased with the initial ramp of Turing and the reduction of inventory in the channel. During the quarter, we filled out our Turing lineup with the launch of midrange GeForce products that enable us to delight gamers with the best performance at every price point starting at \$149.

New product launches this quarter included the GeForce GTX 1660 Ti, 1660, and 1650, which bring Turing to the high-volume PC gaming segment for both desktop and laptop. These GPUs deliver up to 50% performance improvement over their Pascal-based predecessors, leveraging new share innovations such as concurrent floating point and integer operations, a unified cache and Adaptive Shading, all with the incredibly power efficient architecture.

We expect continued growth in the gaming laptops this year. GeForce gaming laptops are one of the shining spots of the consumer PC market. This year, OEMs have built a record of nearly 100 GeForce gaming laptops. GeForce laptops start at \$799 and all the way up to an amazing GeForce RTX 2080 4K laptops that are more powerful than even next-generation consoles. The content ecosystem for ray-traced games is gaining significant momentum.

At the March Game Developers Conference, ray-tracing sessions were packed. Support for ray-tracing was announced by the industry's most important game engines, Microsoft DXR, Epic's Unreal Engine and Unity. Ray-tracing will be the standard for next-generation games.

In March, at our GPU Technology Conference, we also announced more details on our cloud gaming strategies through our GeForce NOW service and the newly announced

GFN Alliance. GeForce NOW is a GeForce gaming PC in the cloud, for the 1 billion PCs that are not game ready, expanding our reach well beyond today's 200 million GeForce gamers. It's an open platform that allows gamers to play the games they own instantly in the cloud, on any PC or Mac anywhere they like. The service currently has 300,000 monthly active users with 1 million more on the waitlist.

To scale out to millions of gamers worldwide, we announced the GeForce NOW Alliance, expanding GFN through partnerships with the global telecom providers. SoftBank in Japan and LG Uplus in South Korea will be among the first to launch GFN later this year.

NVIDIA will develop the software and manage the service and share the subscription revenue with alliance partners. GFN runs on NVIDIA's edge computing servers. As Telcos race to offer the new services for their 5G network, GFN is an ideal new 5G application.

Moving to data center. Revenue was \$634 million, down 10% year-on-year and down 7% sequentially, reflecting the pause in hyperscale spending. While demand from some hyperscale customers bounced back nicely, others paused or cut back. Despite the uneven demand back drop -- backdrop, the quarter had significant positives, consistent with the growth drivers we outlined on our previous earnings call.

First, inference revenue was up sharply both year-on-year and sequentially with broad-based adoption across a number of hyperscale and consumer Internet companies. As announced at GTC Amazon and Alibaba joined other hyperscalers, such as Google, Baidu and Tencent in adopting the T4 in their data centers.

A growing list of consumer Internet companies is also adapting our GPUs for inference, including LinkedIn, Expedia, Microsoft, PayPal, Pinterest, Snap and Twitter. The contribution of inference to our data center revenue is now well into the double-digit percent.

Second, we expanded our reach in enterprise, teaming up with major OEMs to introduce the T4 enterprise and edge computing servers. These are optimized to run the NVIDIA, CUDA-X AI acceleration libraries for AI and data analytics. With an easy to deploy software stack from NVIDIA and our ecosystem partners, this wave of NVIDIA edge AI computing systems enables companies in the world's largest industries, transportation, manufacturing, industrial, retail, health care and agricultural, to bring intelligence to the edge where the customers operate.

And third, we made significant progress in data center rendering and graphics. We unveiled a new RTX Server configuration packing 40 GPUs into an 8 use case and up to 32 servers in a pod, providing unparalleled density, efficiency and scalability. With a complete stack, this server design is optimized for three data center graphic workloads: rendering, remote workstations and cloud gaming.

The rendering opportunity is starting to take shape with early RTX Server deployments at leading studios, including Disney, Pixar and Weta. In the quarter, we announced our pending acquisition of Mellanox for \$125 per share in cash, representing a total enterprise

value of approximately \$6.9 billion, which we believe will strengthen our strategic position in data center. Once complete, the acquisition will unite two of the world's leading companies in high-performance computing. Together, NVIDIA's computing platform and Mellanox's interconnects power over 250 of the world's TOP500 supercomputer and have as customers every major cloud service provider and computer maker.

Data centers in the future will be architected as giant compute engines with tens of thousands of compute nodes, designed holistically with their interconnects for optimal performance. With Mellanox, NVIDIA will optimize data center scale workloads across the entire computing, networking, and storage stack to achieve higher performance, greater utilization and lower operating cost for customers. Together, we can create better AI computing systems for the cloud to enterprise to the edge. As stated at the time of the announcement, we look forward to closing the acquisition by the end of this calendar year.

Moving to Pro Visualization. Revenue reached \$266 million, up 6% from the prior year and down 9% sequentially. Year-on-year growth was driven by both desktop and mobile workstations, while the sequential decline was largely seasonal. Areas of strength included the public sector, oil and gas and manufacturing. Emerging applications, such as AI, AR, VR, contributed an estimated 38% of pro visualization revenue. The real time ray-tracing capabilities of RTX are a game changer for the visual effects industry, and we are seeing tremendous momentum in the ecosystem.

At GTC, we announced that the world's top 3D application providers have adopted NVIDIA RTX in their product releases set for later this year, including Adobe, Autodesk, Chaos Group, Daassault and Pixar. With this rich software ecosystem, NVIDIA RTX is transforming the 3D market. For example, Pixar is using NVIDIA RTX ray-tracing on its upcoming films. Weta Digital is using it for upcoming Disney projects and Siemens NX ray-traced studios users will be able to generate rendered images up to four times faster in their product design workloads. We are excited to see the tremendous value that NVIDIA RTX is bringing to the millions of creators and designers served by ecosystem partners.

Finally, turning to automotive. Q1 revenue was \$166 million, up 14% from a year ago and up 2% sequentially. Year-on-year growth was driven by growing adoption of next-generation AI cockpit solutions and autonomous vehicle development deals. At GTC, we had major customer and product announcements.

Toyota selected NVIDIA's end-to-end platform to develop, train and validate self-driving vehicles. This broad partnership includes advancements in AI computing, infrastructure using NVIDIA GPUs, simulation using NVIDIA DRIVE Constellation platform and in-car AV computers based on the DRIVE, AGX, Xavier or Pegasus. We also announced the public availability of DRIVE Constellation, which enables millions of miles to be driven in virtual worlds across the broad range of scenarios with greater efficiency, cost effectiveness and safety than what's possible to achieve in the real world. Constellation will be reported in our data center market platform.

And we introduced NVIDIA Safety Force Field, a computational defensive driving framework that shields autonomous vehicles from collisions. Mathematically verified and validated in simulation, Safety Force Field will prevent a vehicle from creating, escalating or contributing to an unsafe driving situation.

We continue to believe that every vehicle will have an autonomous capability one day, whether with a driver or driverless. To help make that vision a reality, NVIDIA has created an end-to-end platform for autonomous vehicles from AI computing infrastructure to simulation to in-car computing. And Toyota is our first major win that validates this strategy. We see this as a \$30 billion addressable market by 2025.

Moving to the rest of the P&L and balance sheet. Q1 GAAP gross margin was 58.4% and non-GAAP was 59%, down year-on-year to lower gaming margins and mix, up sequentially from Q4, which had a \$128 million charge from DRAM boards and other components. GAAP operating expenses were \$938 million, and non-GAAP operating expenses were \$753 million, up 21% and 16% year-on-year, respectively. We remain on track for high single-digit OpEx growth in fiscal 2020 while continuing to invest in the key platforms driving our long-term growth, namely graphics, AI and self-driving cars.

GAAP EPS was \$0.64, and non-GAAP EPS was \$0.88. We did not make any stock repurchases in the quarter. Following the announcement of the pending Mellanox acquisition, we remain committed to returning \$3 billion to shareholders through the end of fiscal 2020 in the form of dividends and repurchases. So far, we have returned \$800 million through share repurchases and quarterly cash dividends.

With that, let me turn to the outlook for the second quarter of fiscal 2020. While we anticipate substantial quarter-over-quarter growth, our Q2 outlook is somewhat lower than our expectations earlier in the quarter where our outlook for fiscal 2020 revenue was flat to down slightly from fiscal 2019. The data center spending pause around the world, will likely persist in the second quarter and visibility remains low. In gaming, the CPU shortages while improving will affect the initial ramp of our laptop business.

For Q2, we expect revenue to be \$2.55 billion, plus or minus 2%. We expect a stronger second half than and first half, and we are returning to our practice of providing revenue outlook one quarter at a time. Q2 GAAP and non-GAAP gross margins are expected to be 59.2% and 59.5%, respectively, plus or minus 50 basis points. GAAP and non-GAAP operating expenses are expected to be approximately \$985 million, \$765 million, respectively.

GAAP and non-GAAP OI&E are both expected to be income of approximately \$27 million. GAAP and non-GAAP tax rates are both expected to be 10%, plus or minus 1%, excluding discrete items. Capital expenditures are expected to be approximately \$120 million to \$140 million. Further financial details are included in the CFO commentary and other information available on our IR website.

In closing, let me highlight upcoming events for the financial community. We'll be presenting at the Bank of America Global Technology Conference on June 5th, at the RBC

Future of Mobility Conference on June 6th, and at the NASDAQ investor conference on June 13th. Our next earnings call to discuss financial results for the second quarter of fiscal 2020 will take place on August 15th.

We'll now open the call for questions. Operator, will you please poll for questions? Thank you.

Questions And Answers

Operator

(Question And Answer)

(Operator Instructions). And your first question comes from the line of Aaron Rakers with Wells Fargo.

Q - Aaron Rakers {BIO 6649630 <GO>}

Yes, thanks for taking the question. Colette, I was wondering if you could give a little bit more color or discussion around what exactly you've seen in the data center segment and whether or not or what you're looking for in terms of signs that we can kind of return to growth or maybe this pause is behind us. I guess what I'm really asking is kind of what's changed over the last, let's call it three months relative to your prior commentary from a visibility perspective and just demand perspective within that segment.

A - Colette Kress {BIO 18297352 <GO>}

Sure. Thanks for the question as we start out here. I think when we had discussed our overall data center business three months ago, we did indicate that our visibility as we turned into the new calendar year was low. We had a challenge in terms of completing some of the deals at the end of that quarter.

As we moved into Q1, I think we felt solid in terms of how we completed. We saw probably a combination of those moving forward continuing with their CapEx expenditures and building out in terms of what they need for the data centers. Some others are still in terms of the pause. So as we look in terms of with Q2. I think we see a continuation of what we have in terms of the visibility, not the best visibility going forward, but still rock-solid to what we think are benefits of what we provide in terms of a platform. Our overall priorities are aligned to what we see with the hyperscalers as well as the enterprises, as they think about using AI in so many of their different workloads. But we'll just have to see as we go forward, how this turns out. But right now, visibility probably just remains the same about as where we were when we started three months ago.

Q - Aaron Rakers {BIO 6649630 <GO>}

Okay. And then as a quick follow-up on the gaming side, last quarter you talked about that being down -- I think it was termed as being down slightly for the full-year. Is that still the expectation? Or how has that changed?

A - Colette Kress {BIO 18297352 <GO>}

So at this time, we don't plan on giving a full year overall guidance. I think our look in terms of gaming, all of the drivers that we thought about earlier in the quarter and we talked about on our Investor Day and we have continued to talk about are still definitely in line. The drivers of our gaming business and Turing, RTX for the future are still on track, but we're not providing guidance at this time for the full-year.

Operator

And your next question comes from the line of Harlan Sur of JP Morgan.

Q - Harlan Sur {BIO 6539622 <GO>}

Good afternoon. Thanks for taking my question. On the last earnings call, you had mentioned, China gaming demand as a headwind. At the Analyst Day in mid-March, I think Jensen had mentioned that the team was already starting to see better demand trends out of China may be given the relaxed stance on gaming bans. Do you anticipate continued China gaming demand on a go-forward basis? And maybe talk about some of the dynamics driving that demand profile in the China geography.

A - Jen-Hsun Huang

Sure, China looks fine. I think China has stabilized. The gaming market in China is really vibrant and it continues to be vibrant. Tencent's releasing new games. I think you might have heard that Epic stores now are open in Asia, and games are available from the West. So there are all kinds of positive signs in China. There are some 300 million PC gamers in China, and people are expecting it to grow. We're expecting the total number of gamers to continue to grow from the 1 plus billion PC gamers around the world to something more than that. And so thing look fine.

Q - Harlan Sur {BIO 6539622 <GO>}

Thanks for that. And then, as a follow-up, a big part of the demand profile in the second half of the year for the gaming business is always the lineup of AAA rated games. Obviously, you guys have a very close partnership with all of the game developers. How does the pipeline of new games look? Kind of they get launched October, November time frame, either total number of blockbuster games and also games supporting real-time ray-tracing as well as some of your DLSS capabilities.

A - Jen-Hsun Huang

Yes. Well, it's seasonal. Second half of the year, we expect to see some great games. We won't preannounce anybody else's games for them, but this is a great PC cycle, because it's the end of the console cycle, and PCs is just where the action's at these days with battle royale and eSports, and so much social going on, the PC gaming ecosystem is just really vibrant.

Our strategy with RTX was to take a lead and move the world to ray-tracing. And at this point, it's fairly safe to say that the leadership position that we've taken has turning into a

movement that has turned next-generation gaming ray-tracing into a standard. Almost every single game platform will have to have ray-tracing and some of them already announced it.

And the partnerships that we've developed are fantastic, Microsoft DXR is supporting ray-tracing, Unity is supporting ray-tracing, Epic is supporting ray-tracing. Leading publishers like EA has adopted RTX and supporting ray-tracing. And movie studios, Pixar has adopted -- announced that they're using RTX and will use RTX to accelerate their rendering of films. And so Adobe and Autodesk jumped on top -- jumped onto RTX and will bring ray-tracing to their content and their tools. And so, I think at this point, it's fair to say that ray-tracing is the next-generation and is going to be adopted all over the world.

Operator

And your next question comes from the line of Timothy Arcuri with UBS

Q - Timothy Arcuri {BIO 3824613 <GO>}

Thank you. I guess the first question is for Colette. So what went into the decision to pull full year guidance versus just cutting it? Is it really to see around how long it could take for data center to come back? Thank you.

A - Colette Kress {BIO 18297352 <GO>}

Yes. I'll start off here and kind of go back to where our thoughts were in Q1 and why we provided full year guidance when we were in Q1. When we looked at Q1 and what we were guiding, we understood that it was certainly an extraordinary quarter, something that we didn't feel was a true representative of our business, and we wanted to get a better view of our trajectory of our business in terms of going forward. We are still experiencing. I think the uncertainty as a result of the pause in terms of with the overall hyperscale data centers. And we do believe that's going to extend into Q2.

However, we do know and expect that our Q2 -- or excuse me, our H2 will likely be sizably larger than our overall H1, and the core dynamics of our business at every level is exactly what we expected. Just that said though, we're going to return to just quarterly guidance at this time.

Q - Timothy Arcuri {BIO 3824613 <GO>}

Okay. Thanks. And then just as a follow-up, can you give us some even qualitative, if not quantitative, sense of the \$320 million incremental revenue for July, how that breaks out? Is the thinking sort of that data center is going to be flat to maybe up a little bit and pretty much the remainder of the growth comes from gaming? Thanks

A - Colette Kress {BIO 18297352 <GO>}

Yes. So, when you think about our growth between Q1 and Q2, yes, we do expect in terms of our gaming to increase. We do expect our Nintendo Switch to switch to start again in

sizable amounts once we move into Q2, and we do at this time expect probably our data center business to grow.

Operator

And your next question comes from the line of Toshiya Hari with Goldman Sachs.

Q - Toshiya Hari {BIO 6770302 <GO>}

Thanks for taking the question. Jensen, I had a follow-up on the data center business. I was hoping you could provide some color in terms of what you're seeing not only from your hyperscale customers, which you've talked extensively on, but more on the enterprise and HP side of your business and specifically on the hyperscale side. You guys talked about this pause that you're seeing from your customer base. When you're having conversations with your customers, do they give you a reason as to why they're pausing? Is it too much inventory of GPUs and CPUs and so on and so forth? Or is it optimization giving them extra capacity? Is it caution on their own business going forward? Or is it a combination of all the above? Any color on that would be helpful too. Thank you.

A - Jen-Hsun Huang

Hyperscalars are digesting the capacity they have. At this point, I think it's fairly clear that in the second half of last year they took on a little bit too much capacity. And so everybody has paused to give themselves a chance to digest. However, our business on inference is doing great, and we're working with CSPs all over the world to accelerate their inference models. Now, the reason why recently the inference activity has gotten just off the charts is because of breakthroughs in what we call conversational AI. In fact today, I think that -- I just saw it today, but I've known about this work for some time.

Harry Shum's Group, Microsoft AI Research Group, today announced their multi-task DNN general language understanding model. And it broke benchmark records all over the place. And basically, what this means is the three fundamental components of conversational AI, which is speech recognition; natural language understanding which does multi-task DNN is a breakthrough, and it's based on a piece of work that Google did recently called BERT.

And then, text to speech. All of the major pieces of a conversational AI are now put together. Of course, it's going to continue to evolve, but these models are gigantic to train. And in the case of Microsoft's network, it was trained Volta GPUs, and these systems require large amounts of memory. The models are enormous. It takes an enormous amount of time to train these systems.

And so we're seeing a breakthrough in conversational AI. And across the board internet companies would like to make their AI much more conversational, so that you can access your phones and smart speakers and be able to engage AI practically everywhere. The work that we're doing in the industries makes a ton of sense. We're seeing AI adoption in just about all the industries from transportation to Health care to retail, to logistics, industrials, agriculture. And the reason for that is because they have a vast amount of data that they are collecting and I heard a statistics just the other day from a talk that Satya

gave that some 90% of today's data was created just two years ago and is being created by and gathered by these industrial systems all over the world.

And so if you want to put that data to work -- and you could create the models using our systems, our GPUs for training and then you can extend that all the way out to the edge. This last quarter, we started to talk about our enterprise server based on T4. This Inference Engine that has been really successful for us at the CSPs is now going out into the edge and we call them edge servers and enterprise servers. And these edge systems are going to do AI, basically instantaneously. It's too much data to move all the way to the cloud. You might have data sovereignty concerns. You want to have very, very low latency. Maybe it needs to have multi-sensor fusion capabilities, so it understands the context better.

For example, what it sees and what it hears has to be harmonious. And so you need that kind of AI, those kind of sensor computing at the edge. And so we're seeing a ton of excitement around this area. Some people call it intelligent edge, some people call it edge computing and now with 5G networks coming, we're seeing a lot of interest around the edge computing servers that we are making. And so those are the activities that we're seeing.

Q - Toshiya Hari {BIO 6770302 <GO>}

Thank you. And as a quick follow-up on the gaming site. Colette, can you characterize product mix within gaming? You saw in the current quarter -- you cited mix is one of the key reasons why gross margins were down year-over-year, albeit off a high base. Going into Q2 in the back half, would you expect SKU mix within gaming to improve or stay the same? I asked because it's important for gross margins obviously. Thank you.

A - Colette Kress {BIO 18297352 <GO>}

Yes. When you look at our sequential gross margin increase, that will be influenced by our larger revenue -- our larger revenue and better mix, which you are correct, is our largest driver of our overall gross margin. However, we will be beginning the Nintendo Switch backup, and that does have lower gross margins than the company average, influencing therefor Q2 gross margin guidance that we had provided.

As we look forward towards the rest of the year, we think mix and the higher revenue again will influence and likely rise our overall gross margins for the full year.

Operator

And your next question comes from the line of Joe Moore with Morgan Stanley.

Q - Joe Moore {BIO 17644779 <GO>}

Great, thank you. You talked quite a bit about GeForce now in the prepared remarks and at the Analyst Day. It seems like cloud gaming is going to be a big topic at E3. Is that going to be your preferred way to go to market with cloud gaming? And do you expect to sell GPUs to sort of traditional cloud vendors in non-GeForce now fashion?

A - Jen-Hsun Huang

Yes. Our strategy for cloud gaming is to extend our PC position for GeForce gamers into the cloud, and our strategy for building out our network is partnerships with Telcos around the world. And so, we'll build out some of it. And on top of the service, we have our entire PC gaming stack. And when we host the service we'll move to a subscription model. And with our Telcos around the world who would like to provision the service at their edge servers and many of them would like to do so in conjunction with our 5G Telco services to offer cloud gaming as a differentiator. And all of these different countries where PC exposure has been relatively low, we have an opportunity to extend our platform out to that billion PC gamers. And so that's our basic strategy.

We also offer our edge server platform to all of the cloud service providers. Google has NVIDIA GPU graphics in the cloud, Amazon has NVIDIA GPU graphics in the cloud and Microsoft has NVIDIA's GPU graphics in the cloud. And these GPUs will be fantastic also for cloud gaming and workstation graphics and also ray-tracing. And so the platform is capable of running all of the things that NVIDIA runs, and we try to put it in every data center, in every cloud from every region that's possible.

Q - Joe Moore {BIO 17644779 <GO>}

Thank you very much.

Operator

And your next question comes from the line of Vivek Arya with Bank of America Merrill Lynch.

Q - Vivek Arya {BIO 6781604 <GO>}

Thanks for taking my question. Actually, I had a clarification for Colette and a question for Jensen. Colette, are you now satisfied that the PC gaming business is operating at a normal level when you look at the Q2 guidance? Like are all the issues regarding inventory and other issues, are they over? Or do you think that the second half of the year is more the normalized run rate for your PC gaming business?

And then Jensen, on the data center, NVIDIA has dominated the training market. Inference sounds a lot more fragmented and competitive. There's a lot of talk of software being like more at the framework level. How should we get the confidence that your lead in training will help you maintain a good lead in inference also? Thank you.

A - Colette Kress {BIO 18297352 <GO>}

Thanks for the question. So let's start with your first part of the question regarding how we reach overall normalized gaming levels. When we look at our overall inventory in the channel, we believe that this is relatively behind us and moving forward that it will not be an issue. Going forward, we will probably reach normalized level for gaming somewhere between Q2 and Q3, similar to our discussion that we had back at Analyst Day, at the beginning of the quarter.

A - Jen-Hsun Huang

NVIDIA's strategy is accelerated computing. It is very different than an accelerator strategy. For example, if you were building a smart microphone, you need an accelerator for speech recognition, ASR. Our company is focused on accelerated computing, and the reason for that is because the world's body of software is really gigantic, and the world's body of software continues to evolve and AI is nowhere near done. We're probably at the first or a couple of innings of AI at that. And so the amount of the software and the size of the models are going to have to continue to evolve.

Our accelerated computing platform is designed to enable the computer industry to bring forward into the future all the software that exists today, whether it's TensorFlow or Caffe or PyTorch or it could be classical machine learning algorithms like XGBoost, which is actually right now the most popular framework and machine learning overall. And there are so many different types of classical algorithms and not to mention all of the handwritten engineered algorithms by programmers. And those algorithms and those hand engineered algorithms also would like to be mixed in with all of the deep learning or otherwise classical machine learning algorithms.

This whole body of software doesn't run on a single function accelerator. If you would like that body of software to run on something, it would have to be sufficiently general purpose. And so the balance that we made was we invented this thing called a Tensor Core that allows us to accelerate deep learning to the speed of light. Meanwhile, it has the flexibility of CUDA so that we can bring forward everything in classical machine learning as people have started to see with RAPIDS and is being announced being integrated into machine learning pipelines in the cloud and elsewhere, and then also all of the high-performance computing applications or computer vision algorithms, image processing algorithms that don't have deep learning or machine learning alternatives.

And so our company's focused on accelerated computing. And speaking of inference, that's one of the reasons why we're so successful in inference right now. We're seeing really great pickup. And the reason for that is because the type of models that people want to run on one application, and let's just use one application, one very-very exciting one conversational AI, you would have to do speech recognition. You would have to then do natural language understanding to understand what the speech is. You might have to convert, you have to translate to another language. Then you have to do something related to maybe making a recommendation or making a search. And then after that, you have to convert that recommendation and search and the intent into speech.

While some of it could be Eight-bit integers, some of it really wants to be 16 bit floating point and some of it because of the development state of it may want to be in 32-bit floating point. And so the mixed precision nature and the computational algorithm nature or flexibility nature of our approach make it possible for cloud providers and people who are developing AI applications, to not have to worry about exactly what model it runs or not. We run every single model. And if it doesn't currently run, we'll help you make it run.

And so the flexibility of our architecture and the incredible performance in deep learning, is really a great balance, and allows customers to deploy it easily. So, our strategy is very

different than an accelerator. I think the only accelerators that I really see successful at the moment are the ones that go into smart speakers. And surely, there are a whole bunch being talked about, but I think the real challenge is, how to make it run real workloads. And we're going to keep cranking along in our current strategy, and keep raising the bar as we have in the past.

Q - Vivek Arya {BIO 6781604 <GO>}

Thank you.

Operator

Your next question comes from the line of Stacy Rasgon with Bernstein Research.

Q - Stacy Rasgon {BIO 16423886 <GO>}

Hi guys, thanks for taking my question. And this is a question for Colette. Colette, so you said inference and rendering within data center were both up very strongly. But I guess that has to imply that like the training/acceleration pieces is quite weak even weaker than the overall. And given those should be adding to efficiency. I'm just surprised it's down that much, I mean is this truly just digestion? I mean is it share? I mean like your competitor is now shipping some parts here. I mean I guess how do we get confidence that just we haven't seen the ceiling on this? I mean do you think given the trajectory, you can exit the year above the prior peaks? I guess you kind of have to give at least the qualitative outlook in the second -- I guess maybe just any color you can give us on any of those trends would be super helpful.

A - Colette Kress {BIO 18297352 <GO>}

As we discussed, Stacy, we are seeing many of the hyperscalers definitely purchasing in terms of the Inferencing into the installment that it continues, also in terms of the training instances that they will need for their cloud or for internal use absolutely. But we have some that have paused and are going through those periods.

So that -- we do believe this will come back. We do believe as we look out into the future that they will need that overall deep learning for much of their research as well as many of their workloads.

So no concern on that, but right now, we do see a pause. I'll turn it over to Jensen to see if he has any additional comments.

A - Jen-Hsun Huang

Let's see. I think that when it comes down to training, if your infrastructure team tells you not to buy anything, the thing that suffers is time to market and some amount of experimentation that allows you to get better comps and waiting longer. And I think that for computer vision type of algorithms and recommendation type of algorithms, those -- that posture may not be impossible.

However, the type of work that everybody is now jumping on top of which is natural language understanding and conversational AI and the breakthrough that Microsoft just announced, if you want to keep up with that, you're going to have to buy much, much larger machines. And I'm looking forward to that, and I expect that that's going to happen. But in the latter part of last year, Q4 and Q1 of this year, we did see pause from the hyperscalers and -- but I don't expect it to last.

Q - Stacy Rasgon {BIO 16423886 <GO>}

Got it. Just as a quick follow-up, I just wanted to ask about the regulatory around Mellanox in the context of what we're seeing out of China now. How do we sort of gauge the risk of, I guess, potential for the deterioration in relationships sort of spilling over on the regulatory front around that deal? We've seen that obviously with some of the other large deals in the space. What are your thoughts on that?

A - Jen-Hsun Huang

Well, on first principles, the acquisition is going to enable data centers around the world, whether it's U.S. Or elsewhere in China, to be able to advance much-much more quickly. We're going to invest and building infrastructure technology. And as a combined company, we'll be able to do that much better. And so this is good for customers, and it's great for customers in China.

The two matters, whether it's -- the two matters that we're talking about are just -- are different. And one is related to competition in a -- with respect to our acquisition, the competition in the market, and the other one is related to trade. And so the two matters are just different. And in our particular case, we bring so much value to the marketplace in China, and I'm confident that the market will see that.

Operator

And your next question comes from the line of C.J. Muse with Evercore ISI.

Q - CJ Muse

Yes. Good afternoon. Thank you for taking my question. I guess a question on the non-cloud part of your data center business. So if you think about the trends you're seeing in enterprise virtualization, in HPC and all the work you're doing around RAPIDS, rendering etcetera, can you kind of talk through the visibility you have today for that part of your business? I think that's roughly 50% of the mix. So is that a piece that you feel confident you can grow in 2019? And any color around that would be appreciated.

A - Jen-Hsun Huang

We expect it to grow in 2019. A lot of our T4 inference work is related to what people call edge computing. And it has to be done at the edge because of the amount of data that otherwise would be transferred to the cloud is just too much. It has to be done at the edge because of data sovereignty issues and data privacy issues. And it has to be done at the edge because the latency requirement is really, really high. It has to respond basically like a reflex and to make a prediction or make a suggestion or stop a piece of machinery

instantaneously. And so a lot of that work, a lot of the work that we're doing in T4 inference is partly in the cloud, a lot of it is at the edge.

T4 servers for enterprise were announced, I guess about halfway through the quarter and the OEMs are super excited about that because the number of companies in the world who want to do data analytics, predictive analytics is quite large, and the size of the data is growing so significantly. And with Moore's Law ending, it's really hard to power through terabytes of data at a time.

And so we've been working on building the software stack from the new memory architectures and storage architectures all the way to the computational middleware, and it's called RAPIDS and I appreciate you saying that, and that's being put together. And the activity in GitHub is just fantastic.

And so you can see all kinds of companies jumping in to make contributions, because they would like to be able to take that open source software, and run it in their own data center on our GPUs. And so I expect the enterprise side of our business, both for enterprise big data analytics or for edge computing to be a really good growth driver for us this year.

Q - CJ Muse

As follow-up, real quickly on auto. It's a business that you talked about, more R&D focused, but clearly I think has surprised positively. What's the visibility like there? And how should we think about growth trajectory into the second half of the year?

A - Jen-Hsun Huang

Our automotive strategy has several components. There's the engineering component of it where we -- our engineers and their engineers have to co-develop the autonomous vehicles. And then there's three other components. There's the component of AI computing infrastructure we call DGX and/or any of the OEM service that include our GPUs that are used for developing the AIs. The cars are collecting a couple of terabytes per day, per test car, and all of that data has to be powered through and crunched through in the data center. And so we have an infrastructure of what we call DGX that people could use and we're seeing a lot of success there.

We just announced this last quarter, a new infrastructure called Constellation that lets you essentially drive thousands-and-thousands of test cars in your data center, and they're all going through pseudo directed random or directed scenarios that allows you to either test untestable scenarios, or regress against previous scenarios and we call that Constellation.

And then lastly, after working on a car for several years, we would install the computer inside the car and we call that DRIVE And so these are the four components of opportunities that we have in the automotive industry. We're doing great in China. There's a whole bunch of electric vehicles being created. The robotaxis' developments around the world largely use NVIDIA's technology. We recently announced the partnership with

Toyota. There's a whole bunch of stuff that we're working on, and I'm anxious to announce them to you.

But this is an area that is the tip of the iceberg of a larger space we call robotics and computing at the edge. But if you think about the basic computational pipeline of a self-driving car, it's no different essentially than a smart retail or the future of computational medical instruments, agriculture, industrial inspection, delivery drones, all basically use essentially the same technique. And so this is the foundational work that we're going to do for a larger space that people call the intelligent edge or computing at the edge.

Operator

Your next question comes from the line of Chris Caso with Raymond James.

Q - Chris Castle {BIO 6120611 <GO>}

Yes. Thank you. Good afternoon. First question is on notebooks and just to clarify what's been different from your expectations this year. Is it simply that the OEMs didn't launch the new models you'd expected given the shortage? Or is it more just about unit volume?

And then just following up on that, what's your level of confidence in that coming back to be a driver as you go into the second half of the year?

A - Jen-Hsun Huang

In Q2, we were -- we had to deal with some CPU shortage issues at the OEMs. It's improving, but the initial ramp will be affected. And so the CPU shortage situation has been described fairly broadly, and it affected our initial ramp. We don't expect it to affect our ramp going forward. And the new category of gaming notebooks that we created called Max-Q has made it possible for really amazing gaming performance to fit into a thin and light.

And these new generations of notebooks with our Max-Q design and the Turing GPU, which is super energy efficient, in combination, made it possible for OEMs to create notebooks that are both affordable, all the way down to \$799, thin, and really delightful, all the way up to something incredible with RTX 2080 and a 4K display. And these are thin notebooks that are really beautiful that people would love to use.

And this -- the invention of the Max-Q design method and all the software that went into it that we announced last year, we had -- I think last year, we had some 40 notebooks or so, maybe a little bit less than that. And this year, we have some 100 notebooks that are being designed at different price segments by different OEMs across different regions. And so I think this year is going to be quite a successful year for notebooks. And it's also the most successful segment of consumer PCs. It's the fastest-growing segment. It is very largely under penetrated because until Max-Q came along, it wasn't really possible to design a notebook that is both great in performance and experience and also something that a gamer would like to own.

And so finally, we've been able to solve that difficult puzzle, and created powerful gaming machines that are inside a notebook, that's really wonderful to own and carry around. And so this is going to be a really -- this is a fast growing segment, and all the OEMs know it. And that's why they put so much energy into creating all these different types of designs and styles and sizes and shapes. And we have 100 Turning GPU notebooks, gaming PCs ramping right now.

Q - Chris Castle {BIO 6120611 <GO>}

It's very helpful. Thank you. As a follow-up, I just want to follow-up on some of the previous questions on the automotive market. And we've been talking about it for a while. Obviously the design cycles are very long, so you do have some visibility, and I guess the question is, when can we expect an acceleration of auto revenue. Is next year the year?

And then, what will be the driver of that in terms of dollar contribution? I presume some of the Level 2+ things you've been talking about would probably mean most likely they're given the amount of volume there. If you can confirm that and just give some color on expectations for drivers.

A - Jen-Hsun Huang

Yes. Level 2+ call it 2020, late 2021, or like 2022-ish. So that's Level 2 Plus. I would say 2019, very-very early for robotaxi; next year, substantially more volume for robotaxis; 2021, bigger volumes for robotaxis. The ASP differences, the amount of computation you put into a robotaxi because of sensor resolutions, sensor diversity and redundancy, the computational redundancy and the richness of the algorithm, all of it put together, it's probably an order of magnitude plus in computation. And so the economics would reflect that.

And so that robotaxi is kind of like next year, the year after, ramp. And then think of Level 2+ as 2021, 2022. Overall, remember that our economics come from 4 different parts. And so there's the NRE components of it, there's the AI development infrastructure, computing infrastructure part of it, the simulation part of it called Constellation and then the economics of the car. And so we just announced Constellation. The enthusiasm around it is really great. Nobody should ever ship anything they don't simulate. And my expectation is that billions of miles we'll get simulated inside a simulator, long before they'll ship it. And so that's a great opportunity for Constellation.

Operator

And the next question comes from the line of Matt Ramsay with Cowen

Q - Matt Ramsay {BIO 17978411 <GO>}

Thank you very much. Good afternoon. I have a few questions, one for Jensen and one for Colette. I guess Jensen, you've done, you said in many forums that the move down to the new process node at 7 nanometer across the business was not really sufficient to have a platform approach and I agree with that. But maybe you could talk a little bit about your product plans, at least in general terms around a 7 nanometer franchise in the gaming

business and also in your training accelerator program. And I wonder if that might be waiting for some of those products or at least anticipation of those might be the cause of a little bit of a pause here.

And secondly, Colette, maybe you could talk us through your expectations. I understand there's a lack of visibility in certain parts of the business on revenue, but maybe you could talk about OpEx trends through the rest of the year where you might have a little more visibility. Thank you.

A - Jen-Hsun Huang

The entire reason for Q4 and Q1 is attributed to oversupply in the channel of as a result of cryptocurrency. It has nothing to do with Turing. In fact, Turing is off to a faster start than Pascal was, and it continues to be on a faster pace than Pascal was. And so the pause in gaming is now behind us. We're in a growth trajectory in gaming. RTX has -- took the lead on ray-tracing and has now become the standard for next-generation gaming, support from basically every major platform and software provider on the planet.

And our notebook growth is going to be really great, because of the Max-Q design that we invented. And the last couple of quarters would also intersect it with -- overlap with the seasonal slowdown that not sell but build the seasonal builds of the Nintendo Switch, and we're going to go back to the normal build cycle. And as Colette said earlier, somewhere between Q2 and Q3, we'll get back to normal levels for gaming.

And so, we're off to a great start in Turing, and I'm super excited about that. And in the second half of the year, we would have fully ramped up from top to bottom, our Turing architecture spanning everything from 179 to as high performance as you like. And we have the best price, best performance, and best GPU at every single price point. And so I think we're in a pretty good shape.

In terms of process nodes, we tend to design our own process with TSMC. If you look at our process and you measure its energy efficiency, it's off the charts. And in fact, if you take our Turing and you compare it against a 7 nanometer GPU on energy efficiency, its incomparable. In fact, the world's 7 nanometer GPU already exists. And it's easy to go and pull that and compare the performance and energy efficiency against one of our Turing GPUs.

And so, the real focus for our engineering team is to engineer a process that makes sense for us, and to create an architecture that is energy efficient. And the combination of those two things allows us to sustain our leadership position. Otherwise, buying off the shelf process is something that we can surely do, but we want to do much more than that.

A - Colette Kress {BIO 18297352 <GO>}

Okay. And second question regarding OpEx trajectory for the rest of the year, we're still on track to our thoughts on leaving the fiscal year with a year-over-year growth and overall OpEx on a non-GAAP basis in the high single digits. We'll see probably an increase sequentially quarter-to-quarter along there, but our year-over-year growth will start to

decline as we will not be growing at the speed that we did in this last year. But I do believe we're on track to meet that goal.

Operator

And I'll now turn back over to Jen-Hsun for any closing remarks.

A - Jen-Hsun Huang

Thanks, everyone. We're glad to be returning to growth. We are focused on driving three growth strategies. First, RTX ray-tracing. It's now clear that ray-tracing is the future of gaming and digital design, and video RTX is leading the way. With the support of Microsoft, DXR, Epic, Unity, Adobe and Autodesk, game publishers like EA, movie studios like Pixar, industry support has been fantastic.

Second, accelerated computing, and AI computing. The pause in hyperscale spending will pass. Accelerated computing and AI, are the greatest forces in computing today, and NVIDIA is leading these movements. Whether cloud or enterprise or AI at the edge for 5G or industries, NVIDIA's one scalable architecture from cloud to edge is the focal point platform for the industry to build AI upon.

Third, robotics. Some call it an embedded AI, some edge AI or autonomous machines. The same computing architecture is used for self-driving cars, pick-and-place robotics arms, delivery drones and smart retail stores. Everything machine -- every machine that moves or machines that watch other things that move, whether with driver or driverless will have Robotics and AI capabilities.

Our strategy is to create an end-to-end platform that spans NVIDIA's DGX AI computing infrastructure to NVIDIA Constellation simulation to NVIDIA AGX embedded AI computing.

And finally, we're super excited about the pending acquisition of Mellanox. Together, we can advance cloud and edge architectures for HPC and AI computing. See you next quarter.

Operator

And this concludes today's conference call. You may now disconnect.

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 2021, BLOOMBERG LP. All rights reserved. Any reproduction, redistribution or retransmission is expressly prohibited.

FINAL

Bloomberg Transcript