Q2 2018 Earnings Call

Company Participants

- Colette M. Kress, Chief Financial Officer & Executive Vice President
- Jen-Hsun Huang, President, Chief Executive Officer & Director
- Shawn Simmons, Investor Relations

Other Participants

- Atif Malik, Analyst
- Blayne Curtis, Analyst
- C. J. Muse, Analyst
- Chris Caso, Analyst
- Craig A. Ellis, Analyst
- Hans Mosesmann, Analyst
- Joseph L. Moore, Analyst
- Mark Lipacis, Analyst
- Mitch Steves, Analyst
- Robert Mertens, Analyst
- Stacy Aaron Rasgon, Analyst
- Ting Pong Gabriel Ho, Analyst
- Toshiya Hari, Analyst
- Vivek Arya, Analyst
- William Stein, Analyst

MANAGEMENT DISCUSSION SECTION

Operator

Good afternoon. My name is Victoria, and I'll be your conference operator for today. Welcome to NVIDIA's financial results conference call.

Thank you. I will now turn the call over to Shawn Simmons from Investor Relations. Please begin your conference.

Shawn Simmons {BIO 19500814 <GO>}

Thank you. Good afternoon, everyone, and welcome to NVIDIA's conference call for the second quarter of fiscal 2018. 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.

FINAL

Bloomberg Transcript

I'd like to remind you that our call is being webcast live on NVIDIA's Investor Relations website. It's also being recorded. You can hear a replay via telephone until August 17, 2017. The webcast will be available for replay up until next quarter's conference call to discuss Q3 financial results. 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, August 10, 2017, based on information 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 M. Kress {BIO 18297352 <GO>}

Thanks, Shawn. NVIDIA continues to fire on all cylinders. We achieved strong revenue in each of our businesses. Data center revenue grew more than 2.5 times, reflecting momentum behind artificial intelligence as we expanded our product portfolio and began shipping our new Volta platform.

Overall, quarterly revenue reached a record $2.23 billion, up 56% from a year earlier, up 15% sequentially, and well above our outlook of $1.95 billion. From a reporting segment perspective, Q2 GPU revenue increased 59% to $1.9 billion from a year earlier. Tegra processor revenue doubled to $333 million. OEM revenue reached $251 million, reflecting sales of our cryptocurrency-specific GPUs, partially offset by the lapse in our licensing agreement with Intel.

Let's start with our gaming platform. Gaming revenue was $1.19 billion, up 52% year on year and up 15% from Q1. This reflects the vibrant gaming ecosystem, underpinned by continued excitement over our recent launch GPUs and other technologies, great games, and growing interest in e-sports. Gamers continue to love our Pascal-based GPUs, with demand remaining strong for GeForce GTX 10 series products. Our new Max-Q design approach, enabling gaming notebooks that are thinner, lighter, and faster, is finding a strong market. Max-Q is being utilized in more than 20 new notebook models from a wide range of OEMs.

Quality games continue to drive GPU sales. At the E3 Gaming Expo in L.A., we showed the eagerly anticipated Destiny 2 running in 4K on PCs, which drew rave reviews. Major

fall titles in addition to Destiny 2 include the new Call of Duty: WWII, Star Wars: Battlefront 2, and Middle Earth: Shadow of War. We're also seeing the rise of independent titles that have become runaway hits, most notably, a new last man standing shooter called PLAYERUNKNOWN'S BATTLEGROUNDS.

Gaming also continues to be driven by the surging popularity of e-sports. This week's biggest sporting event isn't playing out at Dodger Stadium or Wrigley Field. It's in Seattle at the International Dota 2 Championships. The $24 million in prize money is more than twice the purse of golf's richest event, the U.S. Open. For the third straight year, GeForce GTX is Dota 2's official graphics platform.

GPU sales were lifted by demand from increasingly mining activity, or Ethereum. We serve a large portion of this specialized market with a dedicated board, as seen in our OEM sales, and some with GeForce GTX boards. Our strategy is to stay alert to this fast-changing market, knowing that GPUs are highly efficient at running the algorithms used to mine cryptocurrencies.

Moving to professional visualization, Quadro revenue grew $235 million, up 10% from a year ago, up 15% sequentially, on the demand for high-end real-time rendering and for more powerful mobile workstations. Demand was especially strong in education, both among universities and large public school districts, as well as in the financial sector and defense industry.

And last week at SIGGRAPH's computer graphics show, we highlighted how AI will augment the process of content creation. This new OptiX 5.0 SDK uses AI to accelerate ray tracing when running on our DGX rendering appliance that provides the rendering capability of 150 standard dual-CPU servers. We also introduced NVIDIA eGPU system solutions. This creates a new category for our platforms, using an external chassis to expand access to Titan Xp and Quadro GPUs for the 25 million content creators using standard notebook PCs.

Next, data center, revenue of $416 million was up more than 2.5 times from a year ago. This growth, shared across AI, deep learning, high-performance computing, and GRID, is particularly notable, given that we announced and shipped production units of our Volta-based V100 accelerator as we transition from Pascal generation GPUs. Looking ahead, we see inferencing and video transcoding as emerging applications that are well suited for our GPUs. V100 was among the most important launches at this quarter's GPU Technology Conference [GTC]. It provides 10 times the deep learning power of its year-old predecessor, widely outpacing Moore's Law.

Some of the early V100 production units were given out in recent weeks to leading AI researchers attending the CVPR [Computer Vision and Pattern Recognition] and ICML [International Conference on Machine Learning] conferences. We made available our TensorRT 3 inference optimizer and run time for deep learning application. It delivers 100 times faster inferencing on V100 than the best CPU implementation, and it supports the industry's two most common AI frameworks, Google TensorFlow and Facebook's Caffe.

We also entered into a wide range of important partnerships based on AI. Among them, Baidu has aligned with us on Volta. It's bringing this new architecture to its cloud and optimizing the Paddle Paddle open source deep learning framework for Volta. Volkswagen is collaborating with us to bring the power of AI across their organization. And we announced a new partner program with Taiwan's top ODMs, including Foxconn, Inventec, Quanta, Wistron, to provide them with the early access to the HGX reference architecture, GPU computing technologies, and design guidelines.

Demand remains strong for our DGX AI supercomputer, as organizations take on multiple systems to build out AI-enabled applications. Facebook disclosed a system incorporating 128 DGXs. We have shipped systems to more than 300 unique customers, with 1,000-plus in the pipeline.

Our HPC business remained strong. The new Green500 list showed that the world's 13 most efficient supercomputers run on NVIDIA Tesla P100 accelerators. The top position is held by the Tokyo Institute of Technology's TSUBAME3.0 system. It achieved an extraordinary 14.1 gigaflops per watt, 50% higher efficiency than the previous leader, helping to point the way to exascale supercomputing.

Momentum continued in our GRID graphics virtualization business. Among key wins was Amazon Web Services, whose G3 instances now run on NVIDIA Tesla GPUs.

In automotive, revenue grew to $142 million, up 19% year over year. We announced important new partnerships based on our DRIVE PX AI platform, which is being used by more than 225 car and truck makers, Tier 1 suppliers, HD mapping companies, startups, and research institutions.

Additionally, production cars using our technology continued to make their way into the market. Audi announced last month that the 2019 A8 will be the first car to feature NVIDIA's powered Level 3 autonomous technology. This means that under certain specific use cases, it will be able to be driven by its own software without monitoring from the driver. Our DRIVE PX platform is already on the road in Tesla Motors' full line of cars, including the new Model 3, which uses the second-generation Tesla Autopilot.

This quarter, we also announced that Toyota selected NVIDIA DRIVE PX for their next-generation autonomous cars. Volvo, long a byword for safety and value, and Autoliv selected DRIVE PX for self-driving cars targeted to hit the market by 2021. ZF and HELLA have selected the NVIDIA DRIVE PX platform for their line of autonomous vehicle products. And Baidu announced that its Project Apollo open source self-driving platform for the China market will use NVIDIA DRIVE PX. You'll be hearing more from us in the months ahead at regional GTCs that we will be conducting in Beijing, Munich, Tel Aviv, Taipei, Washington DC, and Tokyo.

Now turning to the rest of the Q2 income statement, GAAP and non-GAAP gross margins for the second quarter were 58.4% and 58.6% respectively, reflecting the decline in Intel licensing revenue. Q2 GAAP operating expenses were $614 million. Non-GAAP operating

expenses were $533 million, up 19% from a year ago, reflecting hiring for our growth initiatives.

GAAP operating income was $688 million. Non-GAAP operating income was $773 million, doubling from a year ago. GAAP net income was $583 million, non-GAAP net income $638 million, also doubling from a year ago, reflecting revenue strength as well as gross margin and operating margin expansion.

For fiscal 2018, we intend to return approximately $1.25 billion to shareholders through stock repurchases and quarterly cash dividends. In the first half of fiscal 2018, we have paid $758 million in share repurchases and $166 million in quarterly cash dividends.

Now turning to the outlook for the third quarter of fiscal 2018, we expect revenue to be $2.35 billion plus or minus 2%. GAAP and non-GAAP gross margins are expected to be 58.6% and 58.8% respectively, plus or minus 50 basis points. GAAP operating expenses are expected to be approximately $672 million. Non-GAAP operating expenses are expected to be approximately $570 million. GAAP OI&E is expected to be an expense of approximately $2 million, inclusive of additional charges from the early conversions of convertible notes. Non-GAAP OI&E is expected to be nominal. GAAP and non-GAAP tax rates are expected to be 17% plus or minus 1%, excluding discrete items.

Further financial details are included in the CFO commentary and other information available on our IR website.

We will now open the call for questions. Please be sure to limit your questions to one or two. Operator, will you please poll for the first question? Thank you.

Q&A

Operator

Certainly. Your first question comes from the line of Vivek Arya with Bank of America.

Q - Vivek Arya {BIO 6781604 <GO>}

Thank you for taking my question, maybe a clarification on a question and a question after that on the data center. On the clarification side, we have seen several quarters where your data center business grew very strongly on a sequential basis. This time the growth was somewhat more modest, and I was wondering if there is a little more color around that.

And then the bigger question is, Jensen, it seems the data center market is bifurcating between your GPU approach on one side and ASICs on the other side. What are you doing to make sure that the balance stays in your favor as the market matures from here?

A - Jen-Hsun Huang {BIO 1782546 <GO>}

FINAL

Bloomberg Transcript

So first of all, Q2 was a transition quarter for our data center. I thought we did great. We almost tripled year over year, and we ramped Volta into volume production. And because Volta was so much better than our last generation processor – Volta is 100 times faster than Kepler, 100 times faster than Kepler just four years ago, and Kepler was already 10 times faster than CPUs. And so Volta was such a giant leap when we announced it in GTC right at the beginning of the quarter. I thought the team did fantastically transitioning the customer base to Volta, and now Volta is in high-volume production.

The application of data center – you asked a larger question about data center. Data center is a very large market, as you know, and the reason for that is because the vast majority of the world's future computing will be largely done in data centers. And there's a very well accepted notion now that GPU acceleration of servers delivers extraordinary value proposition. If you have a data-intensive application, and the vast majority of the future applications in data centers will be data intensive, a GPU could reduce the number of servers you require or increase the amount of throughput pretty substantially. Just adding one GPU to a server could reduce several hundred thousand dollars of reduction in number of servers. And so the value proposition and the cost savings of using GPUs is quite extraordinary.

There are several applications in data centers. First of all, there's training and there's high-performance computing. There's cloud virtual PC, as what Amazon AWS G3 announcement was about this quarter. And then there are also new applications such as inferencing as these models are now going into production, and the new applications that are coming online, which is likely to overwhelm the Internet in the near future, which is live video, consumers taking live video on their phones and sharing it with their friends.

And there are going to be hundreds of millions of these happening all the time, and each one of these videos will have to be transcoded to a variety of formats to be shared with their friends and also has to be – you have to perform AI on it instantaneously so that you could avoid inappropriate video from being streamed to large audiences. And so the number of applications where GPUs are valuable, from training to high-performance computing to virtual PCs to new applications like inferencing and transcoding and AI, are starting to emerge.

The one area where you're talking about ASICs and TPUs, TPU is basically an ASIC. The way to think about that is this. After four generations of evolution of our GPU, NVIDIA GPU is basically a TPU that does a lot more. We could perform deep learning applications, whether it's in training or in inferencing now, starting with the Pascal P4 and the Volta generation. We can inference better than any known ASIC on the market that I've ever seen. And so the new generation of our GPUs is essentially a TPU that does a lot more. And we can do all the things that I just mentioned and the vast number of applications that are emerging in the cloud.

And so our belief is this. Our belief is that, number one, a GPU has to be versatile to handle the vast array of big data and data-intensive applications that are happening in the cloud, because the cloud is a computer. It's not an appliance. It's not a toaster. It's not a lightbulb. It's not a microphone. The cloud has a large number of applications that are data-intensive. And second, we have to be world-class at deep learning, and our GPUs

have evolved into something that can be absolutely world-class at TPU, but it has to do all of the things that a data center needs to do.

Operator

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

Q - Mark Lipacis {BIO 2380059 <GO>}

Thanks for taking my question. It sounds like things went very well on the cryptocurrency side. That market has not had a lot of history, but the little history it has, has had some volatility. And I was wondering if you could help us understand how you think about managing that volatility. And a broader question on this topic is, do you consider cryptocurrency or other blockchain applications on par with your other four big markets?

A - Jen-Hsun Huang {BIO 1782546 <GO>}

Thanks. Cryptocurrency and blockchain is here to stay. The market need for it is going to grow, and over time it will become quite large. It is very clear that new currencies will come to market, and it's very clear that the GPU is just fantastic at cryptography. And as these new algorithms are being developed, the GPU is really quite ideal for it. And so this is a market that is not likely to go away anytime soon, and the only thing that we can probably expect is that there will be more currencies to come. It will come in a whole lot of different nations. It will emerge from time to time, and the GPU is really quite great for it.

What we've done, our strategy is to stay very, very close to the market. We understand its dynamics really well. And we offer the coin miners a special coin-mining SKU. And this product configuration – this GPU configuration is optimized for mining. We stay very close to the market. We know its every single move and we know its dynamics.

And then last thing that I can say is that the larger of a GPU company you are, the greater ability you could absorb the volatility. And so between the combination of the fact that we have GPUs at just about every single price point, we have such incredibly efficient designs that we're so close to the marketplace. And because we have such large volumes, we have the ability to rock and roll with this market as it goes. But this is an important market that likely will continue to grow over time.

Operator

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

Q - Toshiya Hari {BIO 6770302 <GO>}

Great, thanks for taking the question and congrats on the results. I have a question on some of the numbers. So Q2 revenue came in roughly about $250 million above your guide. Can you confirm what some of the drivers were to the upside relative to your guidance? Was it all cryptocurrency, or was it a combination of multiple things? And

related to that, for your Q3 guide I think you are guiding revenue up about $120 million sequentially. What are the puts and takes here on a sequential basis? Thank you.

A - Jen-Hsun Huang {BIO 1782546 <GO>}

Sure, let's see. First of all, we actually gave a really great guidance last quarter, and we beat it by $250 million. And the $250 million you could see in our – we categorized under the OEM SKUs basically the cryptocurrency SKUs. And that, if you reverse-engineered it out, I think is approximately $150 million. And we serve the vast – I would say the large majority of the cryptocurrency demand out of that specialized products.

There are still small miners that buy GeForces here and there, and that probably also increased the demand of GeForces. There were a lot of shortages all over the world. And as we go into this quarter, there's still cryptocurrency mining demand that we know is out there. And based on our analytics and understanding of the marketplace, there will be some amount of demand for the foreseeable future. But it's also the case that there were gamers whose needs and demands were not filled last quarter.

The second quarter is an important part of the year for us. GeForce is in an incredibly great strategic position. After all of the numerous product launches that we've seen from other players, it's very, very clear that the GeForce product lineup is absolutely the best in the world. And the second half is going to see some very exciting titles. You've got Destiny 2. You have Call of Duty from Activision. You have Star Wars: Battlefront from EA. These are going be blockbusters, and we're expecting them to do incredibly well.

We also know that a game that came out of nowhere – and this is one of the things that's really great about the video game market, you never know where the next amazing new title is going to come from. PLAYERUNKNOWN'S BATTLEGROUNDS, it's really essentially Survivor meets Hunger Games. How could that not be a fun game? So they've done incredibly well. And so I think the market dynamic is really vibrant for the second half of the year, and we have a really great position.

With respect to our guidance, the way to think about our guidance, we gave a good guidance and we're comfortable with our guidance. And we know that the dynamics in our business, our data center position is quite exciting. We know that our gaming business is vibrant and our position is excellent. We saw growth across all of our product segments. And we'll just see how it turns out at the end of next quarter.

Operator

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

Q - Stacy Aaron Rasgon {BIO 16423886 <GO>}

Hi, guys. Thanks for taking my questions. First, I was wondering if you could tell us how much Volta contributed to the data center revenue in the quarter. And what are your expectations for that ramp trajectory into the second half? The reason I ask is when I look

at gross margins, they're fine, but it doesn't look like the Volta ramp is driving upside to that number. So I'm trying get some feeling for the trajectory of that ramp.

A - Jen-Hsun Huang {BIO 1782546 <GO>}

First of all, it's very difficult to reverse-engineer from the first ramp of Volta any impact on gross margins, and the reason for that is because the first ramps tend to be more costly, and you're still trying to stabilize yield. There are a lot of complexities involved. But what I can tell you is that we shipped a lot of Voltas. We shipped a lot of Voltas, and Volta is fully ramped. Customers are clamoring for it. The leap generationally for deep learning is quite extraordinary. And so we're expecting Volta to be very, very successful.

Operator

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

Q - C. J. Muse {BIO 6507553 <GO>}

I guess a follow-up question to that on the Volta transition and now that that is ramping in high-volume manufacturing and considering the pretty large uplift in die size there, I'm curious how you're thinking about ASP uplift over time, and whether you would expect that to drive a reacceleration in growth in data center looking into the second half of the calendar year.

A - Jen-Hsun Huang {BIO 1782546 <GO>}

So the first way to think about our ASP is to think about the value proposition that our GPUs provide. Whenever you include a Volta in your data center, in your server that is doing data-intensive processing, the number of commodity servers that it replaces and the number of just NICs [Network Interface Cards] and cables that it replaces is pretty extraordinary. Every single Volta allows you to save several hundred thousand dollars.

And so the price of Volta is driven by the fact that, of course, the manufacturing cost is quite extraordinary. These are expensive things to go and design. The manufacturing cost itself, you guys can estimate it, is probably in the several hundred dollars to close to $1,000. However, the software intensity of developing Volta, the architectural intensity of developing Volta, all of the software intensity associated with all the algorithms and optimizing all the algorithms of Volta is really where the value-add ultimately ends up. And so I guess the pricing – your question relates is pricing. We expect pricing to be quite favorable for Volta.

And then your second question I think is related to acceleration. The data center growth opportunity for us is quite significant, as you know. There are several applications that demand GPUs today. Almost every single data center in the world today recognizes that GPU is the path forward for data-intensive processing. Every single OEM and every single cloud service provider now supports NVIDIA GPUs and offer video GPUs, and Volta is going be the engine that serves them. So I'm expecting a lot of good things from Volta.

Operator

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

Q - Atif Malik {BIO 15866921 <GO>}

Hi, thanks for taking my question, congratulations on the strong results. Even if you exclude the OEM (31:10) contribution, you have beaten the Street expectations. My question is on auto. You've announced a very strong pipeline of auto partnerships this year. Can you just talk about when do you expect acceleration in auto sales? And are there any other ways you can monetize your auto partnerships and maybe through licensing of software stacks? Thank you.

A - Jen-Hsun Huang {BIO 1782546 <GO>}

Sure, thanks a lot, Atif. The roadmap for auto looks like this. For this year and next, what you should see is development partnerships that we have with a growing number of car companies, and they're reflected in NRE projects, development systems, and purchasing of our AI supercomputers like DGX. And so for the next I would say this year and the vast majority of next year, that's what you should expect from the autonomous driving perspective.

Starting next year, you're going to start to see robot taxis start to come to the road. We're working with a handful, maybe I guess about six or seven really exciting robot taxi projects around the world. And you could see them start to go into a prototype or beta testing starting now, and then next year you'll see a lot more of them. And starting 2019, you'll see them go into real commercial services. And so those are robot taxis, what some in the industry call Level 5s, basically driverless cars.

And then the fully autonomous drivered cars, driven cars, branded cars will start hitting the road around 2020 and 2021. So the way to think about it is this year and next is really about development. Starting next year and the following year is robot taxis. And then 2021 to forward you're going to see a lot of Level 4s.

Operator

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

Q - Joseph L. Moore {BIO 17644779 <GO>}

Great, thank you. I wanted to actually ask about the ProVis business. That business continues to grow faster than I had expect it to, and you had real good quarter there. Can you talk about what's driving that and what the trajectory of that business looks like?

A - Jen-Hsun Huang {BIO 1782546 <GO>}

Sure. Our Pro business, call it roughly nearly $1 billion. It grew 8% last year. It grew 8% the year before that, maybe a little bit less, and this year it grew about 10%, maybe a little faster.

The way to think about that business is it's really a platform for design, digital design of all kinds. And it's designing movies, designing cars, designing products, people designing websites. Anybody who's doing digital design could really benefit from a Quadro platform. It's very software-intensive. It's certified with every major computer-aided design package. It's certified by large industrial companies all over the world. You could use Quadro and bring up a database 10 years from now and know that because of the nature of how we manage our software, the certification process we go through with each one of the major industry partners, we could pull up an entire design that was designed five years ago 10 years from now. And so if anything were to happen to a product or a plane or a ship or a building, the level of certainty in your data integrity is going be complete. And so the software intensity is high, and our platform is recognized all over the world as the industry standard.

The growth opportunity for Quadro are several, and it's starting to kick in. And I'm rather optimistic about its future growth as well. One of them is photorealistic rendering. We now have the ability to use our artificial intelligence and ray tracing technology in combination, called OptiX 5.0, that we just announced at SIGGRAPH. That allows you to visualize photorealistic rendering practically interactively, and it's just an amazing thing to watch.

Second, we now have a new system called an external GPU system. That's a partnership between the work that we did with Intel and all of our partners in the ecosystem by taking advantage of Thunderbolt 3 and the new external GPU-capable Windows system. You can now have an external system connect to Thunderbolt and basically our GPU is outside the laptop. And so for some 25 million – 20 million users of thin and light notebooks, you can now have the ability to have a GPU as well and get a boost in your productivity like you've never seen before. And so you can now have thin notebooks and still have the benefit of our GPUs. And so that's a new market for us.

We're going to see virtual reality do quite well, especially in design. And we partnered with HP recently to do an industrial version of a backpack that allows designers to be able to freely roam within their design space and completely in virtual reality. So there's a variety of growth drivers in that business that I'm quite excited about.

Operator

Your next question comes there the line of Craig Ellis with B. Riley.

Q - Craig A. Ellis {BIO 1870408 <GO>}

Thanks for taking the question and congratulations on the very good execution and capitalizing on the crypto opportunity. Jensen, I wanted to start just connecting a few dots. You had mentioned that there was significant upside from that opportunity. And we've seen through checks that we can do via public means that demand was very strong in the quarter. And as I look ahead at the guidance for the fiscal third quarter, it's up 5% when I think normally it would be up in the low double digits. So can you talk about how comfortable you are with your supply availability here, and if demand was there for double-digit growth if you'd be able to achieve that?

And then, Colette, one for you, the OpEx guidance setup, I think it's $37 million quarter on quarter, more than we've seen the last few quarters. Can you just bin out what some of the bigger drivers for that increase are? Thanks, team.

A - Jen-Hsun Huang {BIO 1782546 <GO>}

Sure, thanks a lot. So first of all, to answer that question, I would say there are three factors. The first factor is our strategic position. Our competitive lineup is probably the best it's ever been, better than last year even, which was incredibly strong, better than the year before that because it was incredibly strong. I think our strategic position and the value of our architecture is more powerful today than ever. And so I think number one is our strategic position.

The second, if the demand were there in the second half with respect to – from a perspective of gaming demand and if there's any residual crypto demand, we will surely be able to serve it. And then lastly, the factors related to our guidance, our guidance is we're comfortable with our guidance. We're happy with our guidance, and we want to have an opportunity to come back and give you an update in Q3.

A - Colette M. Kress {BIO 18297352 <GO>}

And, Craig, on our second question regarding the OpEx guidance in Q3, generally our guidance and actuals as we move into Q3 is usually a little stronger, and it's consistent with our normal annual compensation increase that happens in Q3. And also keep in mind, we are expected to move into our new headquarter building within Q3. And underlining our overall growth in investments is our hiring and focus in terms of on AI, autonomous driving, as well as gaming. All of these factors contribute to that with about a 19% year-over-year growth rate in terms of what we're targeting.

Operator

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

Q - Chris Caso {BIO 4815032 <GO>}

Yes, thank you. I just wanted to clarify some earlier comments with regard to Volta and data center. Is it correct to interpret your comments to mean that some customers may have tended to delay purchases as you went through the quarter as they're waiting for Volta given the stronger performance gains for that? And if that's the case, if I've got that right, now that Volta is fully ramped, do you expect that to drive stronger growth rates as you go through the second half?

A - Jen-Hsun Huang {BIO 1782546 <GO>}

The answer to your first question is yes. Volta was a giant leap. It's got 120 teraflops. Another way to think about that is eight of them in one node is essentially one petaflops, which puts it among the top 20 fastest supercomputers on the planet. And the entire world's top 500 supercomputers are only 700 petaflops. And with eight Voltas in one box, we're doing artificial intelligence that represents one of them. So Volta is just a

gigantic leap for deep learning and it's such a gigantic leap for processing that – and we announced it at GTC, if you recall, which is practically right at the beginning of the quarter.

And so the transition was not insignificant, and it was that the team just executed flawlessly. I'm so proud of the team. They executed the most complex processor that's ever been built. And working with our teams, working with our partners at TSMC and Samsung and all of our package partners still, and they just did a great job for us, and so the team did great.

Now looking forward, there's a whole bunch of growth drivers for our data center business. Deep learning is – training is a growth driver. Cloud computing, high-performance computing is a growth driver, and we have new growth drivers with inferencing. And so I'm pretty excited about our prospects going into the age of – the generation of Volta.

In terms of the guidance and what we expect, I think our dynamics are really positive. And so we've just got to – we're happy with the guidance, and let's give you an update at the end of the quarter.

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 and congrats on the very strong results. Jensen, you've had a couple questions already about the pace of growth in data center in the second half from Volta, but I'm thinking a little bit further out. At GTC, you highlighted this $30 billion TAM opportunity by 2020. And when we look at the charts that you've published about your expectation for XFLOPS, the number of XFLOPS required to train an increasing number of deep learning networks through 2020, it looks like your expectation is for that to accelerate over time. But naturally, the Street is contemplating decelerating growth for data center. People don't expect things to grow 150-plus percent forever. So can you comment as to the growth trajectory beyond maybe the very near term in that business? Thank you.

A - Jen-Hsun Huang {BIO 1782546 <GO>}

I think at the highest level, the way to think about that is data-intensive computing, whether it's deep learning or high-performance computing, the GPU is just phenomenal at it. NVIDIA's CUDA GPU was, after 12 years of driving this architecture and pioneering this computing approach, it's just a home run. And the value proposition and the money that it saves people, the amount of energy that it saves is quite extraordinary. One way to think about that is if you speed up an application by a factor of 10, you're basically saying that it takes 10 times fewer servers to do the same job, or you could do 10 times as much work in the same amount of servers. So the value proposition is really quite great.

The applications that we serve is really diverse now. It used to be just high-performance computing and supercomputing. But the number of applications we serve in Internet service providers, manufacturing, healthcare, financial services, transportation, the number of data-intensive applications and industries that need them is really growing very fast.

And so how fast does that – what does that imply in terms of long-term growth? It's hard to say. But first principles would suggest that every single data center in the world will be GPU-accelerated someday. And I've always believed that, and I believe that even more today. Because I believe that in the future, this new computing model that we all finally call AI is going be a highly data-intensive business model, and the GPU is the ideal computing model for that.

So I'm not exactly sure if that completely answers your question, and partly because I'm not exactly sure. I just know that on first principles, the computing architecture is ideal. There's every evidence that every single data center and every single OEM and every single Internet service provider is jumping on this architecture and jumping on Volta. And I believe that AI is going be the future of computing. And so somewhere between those beliefs and executing the business is the truth.

Operator

Your next question comes from the line of Hans Mosesmann with Rosenblatt Securities.

Q - Hans Mosesmann {BIO 1522582 <GO>}

Thanks. Hey, Jensen, can you give us an update in terms of how the new platforms and servers may have impacted the business in the data center with Purley launching here recently and the upcoming EPYC? As a follow-on, when can we expect Volta in the consumer gaming market? Thanks.

A - Jen-Hsun Huang {BIO 1782546 <GO>}

That's a good question, Hans, and it's a good observation. Because Purley, I didn't know if everybody understood that code name, but Purley is a new motherboard, a new platform for Intel servers, and the CPU is Skylake. It's an excellent server platform. And obviously, every OEM and every service provider was waiting for the launch of that, and it officially launched in the middle of this quarter.

And so did it affect the rollout of new servers based on GPUs? It probably did, and surely it did. But now that it's ramped, it's a successful ramp. Every single cloud provider and every single OEM is now fully geared up to take that server to market, and they all have GPU options. Every single OEM in the world now and every cloud provider and every ODM now has NVIDIA GPU chassis and platforms, whether it's in one GPU in 1U to eight GPUs in a supercomputing configuration. And so the number of options of ways to enjoy NVIDIA GPUs is really quite countless now.

Volta for gaming, we haven't announced anything. And all I can say is that our pipeline is filled with some exciting new toys for the gamers, and we have some really exciting new technology to offer them in the pipeline. But for the holiday season for the foreseeable future, I think Pascal is just unbeatable. It's just the best thing out there. And everybody who's looking forward to playing Call of Duty or Destiny 2, if they don't already have one, should run out and get themselves a Pascal.

Operator

Your next question comes from the line of Mitch Steves with RBC Capital Markets.

Q - Mitch Steves {BIO 19155169 <GO>}

Hey, guys. Thanks for taking my question. I just had two high- level ones. First, since you commented on cryptocurrency and blockchain, so when decentralized applications begin to take off and we see people essentially renting out parallel processing, how are you guys going to essentially be able to tell what products are being used in a lease model versus what's being used in gaming, et cetera?

A - Jen-Hsun Huang {BIO 1782546 <GO>}

First of all, it's not really possible because our GPUs are all architecturally compatible, which at some level is one of our strengths. There are hundreds of millions of NVIDIA GPUs in the world, and they're all CUDA compatible, and they're all 100% CUDA compatible. And we're so rigorous and so disciplined about ensuring their compatibility that for developers it's really a wonderful platform.

However, we're thoughtful about how we configure the GPUs so that they're best for the applications. Some applications would like to have the maximum amount of performance in a few nodes. Some would like to have the maximum amount of performance within 30 watts. Some would like to have the maximum amount of flexibility with all of the I/O and connectors and all the display connectors. Some people like to have multi-GPUs and that they have the ability to configure them together. And so every market has a slightly different need, and we have to understand the market needs and understand what it is that the customers are looking for, and configure something that is best for them.

Q - Mitch Steves {BIO 19155169 <GO>}

Got it, thank you, and just one small one on a follow-up. Do you guys have an estimate on how fast the neural network is growing right now relative to a year ago?

A - Jen-Hsun Huang {BIO 1782546 <GO>}

Let's see. A neural net in terms of complexity is approximately – not quite, but approximately doubling every year. And this is one of the exciting things about artificial intelligence. In no time in my history of looking at computers in the last 35 years have we ever seen a double exponential where the GPU computing model, our GPUs are essentially increasing in performance by approximately three times each year. In order to be 100 times in just four years, we have to increase overall system performance by a factor of three, by over a factor of three every year.

And yet on the other hand, on top of it, the neural network architecture and the algorithms that are being developed are improving in accuracy by about twice each year. And so object recognition accuracy is improving by twice each year, or the error rate is decreasing by half each year. And speech recognition is improving by a factor of two each year. And so you've got these two exponentials that are happening, and it's pretty exciting. That's one of the reasons why AI is moving so fast.

Operator

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

Q - Blayne Curtis {BIO 15302785 <GO>}

Thanks for taking my question. On the data center, it was slightly below the Street number. I know it's not your number, but I think we've gotten used to you surpassing the Street number by a wide margin. I'm wondering if you can just talk about in the July quarter the three sub-segments and if they came in as expected. And as you look to October, maybe you can talk about what kind of growth you're looking for, for that segment. Thank you.

A - Jen-Hsun Huang {BIO 1782546 <GO>}

I'm not sure I understand the question. Colette, do you? If you understand, go ahead and answer.

A - Colette M. Kress {BIO 18297352 <GO>}

So in discussing how we did versus our guidance, again, we overshot our overall guidance for Q2. Part of that was the cryptocurrency. And your question was more around the data center and the data center number. We set out with a good amount of work ahead of us to transition to Volta within the quarter. We're very pleased with that result and the overall year-over-year growth that we accomplished in terms of the data center. We always have different ranges across the organization and across the different businesses, but we don't have specifics in terms of our guidance nor did we provide specific guidance in terms of on the data center.

Operator

Your next question comes from the line of Raji Gill with Needham & Company.

Q - Robert Mertens {BIO 20555850 <GO>}

Hi, this is Robert Mertens on behalf of Raji. Congratulations and thank you for taking my question. I guess I wanted to get a little more clarity towards your automotive division. I didn't know if you broke out rough percentages about the near-term growth between infotainment and these AS or autonomous vehicle systems. And then as a follow-up in the autonomous systems, how you're pricing between the different levels, if that's mainly just the GPUs, or if the software is baked in there as well?

FINAL

Bloomberg Transcript

A - Jen-Hsun Huang {BIO 1782546 <GO>}

Our GPUs are useless without software, I'll start from the back and work forward. Our GPUs are useless without software, and the reason for that is because otherwise each one of the markets, whether it's playing games or professional visualization or high-performance computing, doing molecular dynamics computation or doing seismic processing, or perceiving the three-dimensional world around the car and reasoning about where it is and trying to figure out how to drive, all of that software is very, very different.

What we do is there's a core in our company. The core architecture is a GPU core. However, the configuration of the products and the chips and the systems are very different from market to market. So somebody asked me earlier, the gentleman asked me about cryptocurrency. That configuration is very, very different than a gaming configuration, which is different than a high-performance computing configuration, and it's different from our inferencing configuration and it's different from our self-driving car configuration.

And so the chips are designed to be different, even though they're architecturally identical. And then the systems are designed to be market-specific and application-specific. And the software on top of it is super, super application-specific. And that's one of the reasons why our company is increasingly differentiated from a components company and what we call a platform company. Each one of these platforms that we bring to market are very, very different, even though at its core, this data-intensive parallel computing architecture called CUDA is essential among all of them.

We don't break out the automotive from the rest of the Tegra business. The Tegra business consists of basically three parts at the moment. One major component of it is the Nintendo Switch gaming console, and it's just doing incredibly well. I'm so happy for Nintendo because they're risk takers. They're innovators. They're not influenced by what other people do, and they're original thinkers. And I just love the way they invented the Switch and the way they've taken it to market. I'm so happy for them. And it's doing really well.

The second major component is our self-driving car platforms, and a lot of it still is infotainment systems. Our infotainment system is going to evolve into an AI cockpit product line. We initially started with autonomous driving. But you probably heard me say at GTC that our future infotainment systems will basically turn your cockpit or turn your car into an AI. So your whole car will become an AI. It will talk to you. It will know where you are. It knows who's in the cabin. And if there are potential things to be concerned about around the car, it might even just tell you in natural language. And so the entire car will become an AI.

We announced at CES a partnership with Daimler, and they talked about the work that we're doing together in the next-generation car how we're going bring AI to the car. And that's our first visible, highly visible project, and there are others that we're working on. And then the future projects, starting from end of next year with robot taxis, and starting

with 2020 the autonomous cars, fully autonomous cars, you're going to see the rest of that come online, and that's a major component of Tegra.

And then the last component, the third component, major component I would say is AI at the edge. That's the next major revolution. And we have a new product line that we announced about a year ago. It's called Jetson. Jetson is just an amazing little AI computer. And if you want to do deep learning at the edge, whether it's really, really clever cameras for smart and safe cities, to traffic lights that can now monitor traffic, Jetson AI at the edge is going be the next growth opportunity for us. And those three major segments make up essentially the Tegra business. We haven't split each one of them out separately, but one of these days we'll consider doing that.

Operator

Your final question comes from the line of Ambrish Srivastava from BMO.

Q - Ting Pong Gabriel Ho {BIO 2032863 <GO>}

This is Gabriel Ho calling in for Ambrish, thanks for taking my question. I think last year at this time, on the data center business, you disclosed I think a high level of growth for your cloud business. Your deep learning business was about half, and then a third HPC. I was wondering. As you're ramping Volta, how should we think about maybe the mix as you're entering the second half of the year between the PC/cloud and maybe the rest of the business for the data center?

A - Jen-Hsun Huang {BIO 1782546 <GO>}

Ambrish (sic) [Gabriel], that's a good question. Partly I'm not in total control of the answer. But on first principles, let me maybe explain it this way. I believe that there are a few hundred million office workers and information workers whose PCs will be virtualized and just become an application like Netflix, and it will be virtualized. They'll be streamed from our cloud GPUs called GRID.

I believe that every single company in the world, manufacturing, healthcare, finance will use computational approaches to analyzing their business. And some of them will use AI and some of them will use traditional first principle physics-based algorithms. And it's hard to say exactly what the split is going to be. My guess, however, is that AI will be the larger part of that, but you're going to see hybrid versions of it. Most computation – the reason why we're so bullish about CUDA and our GPU, which is able to do both general purpose computation as well as deep learning, is because most algorithms have the combination of both. Inside the card, we don't just use deep learning. We use CUDA and deep learning. We use CUDA for all kinds of algorithms, computer vision algorithms including deep learning. And we're seeing in quantum chemistry, in physics simulations like fluid dynamics, more and more of the algorithms are hybrids of deep learning and numerics. And so that segment of the marketplace is hard to predict.

And then there's just the consumer Internet service providers and the billions and billions of queries that are going into the cloud. Some of them are text, some of them are speech, and increasingly some of them are video. The amount of traffic that's going to be

inferenced using deep learning is going be quite explosive. It's hard to know exactly the pace of each one of these, but I think on first principles, we would all agree that these are large computation challenges, and that the previous model of using just microprocessors to do that computation is not efficient, and that the GPU with its parallel data processing capability and now our fourth-generation deep learning architecture, you essentially have a GPU that does a whole bunch more. And so I think the approach that we take has great promise, and we're just super-enthusiastic about it. But exactly how much it's going contribute in the near term in percentages is going be hard to guess.

Okay, that was great. I appreciate all the questions. We had a great quarter. We're seeing exciting growth dynamics driving in each of our businesses. This is the era of artificial intelligence, and NVIDIA has dedicated ourselves to be its brain. Cloud and Internet service providers are going all in on AI and jumping onto our new Volta GPU. Enterprises and giant industries from transportation to healthcare to manufacturing to financial services have an awakened to the power of AI. And the growing pipeline of the NVIDIA DGX AI supercomputer is a great indicator.

The next revolution of AI will be at the edge, and the most visible impactful evidence will be the autonomous vehicle. Our strategy is to build a ground-up deep learning platform for self-driving cars, and that has put us in pole position to lead the charge. And in gaming, which is actually the first consumer AI application, we have a great strategic position in this growing market. We have a once in a lifetime opportunity ahead. We can make an amazing impact on the future of the world. Thanks for joining us today, and we look forward to giving you another update next quarter.

Operator

This concludes today's conference call. You may now disconnect. Thank you for your participation.