Q1 2022 Earnings Call

Company Participants

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

Other Participants

- Aaron Rakers
- C.J. Muse
- John Pitzer
- Stacy Rasgon
- Timothy Arcuri
- Vivek Arya

Presentation

Operator

Good afternoon. My name is Sinedra, and I will be your conference operator today. At this time, I would like to welcome everyone to the Nvidia's Financial Results Conference Call. All lines have been placed on mute to prevent any background noise. After the speaker's remarks, there will be a question-and-answer session. (Operator Instructions). Thank you.

Simona Jankowski, you may begin your conference.

Simona Jankowski {BIO 7131672 <GO>}

Thank you. Good afternoon, everyone, and welcome to Nvidia's conference call for the first quarter of fiscal 2022. 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 2022. 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.

FINAL

Bloomberg Transcript

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 26, 2021, 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 was exceptionally strong with revenue of $5.66 million and year-on-year growth accelerating to 84%. We set a record in total revenue in gaming, data center and professional visualization driven by our best ever product lineups and structure. Starting with Gaming, revenue of (Technical Difficulty) sequentially and up 100 (Technical Difficulty). This is the third consecutive quarter of accelerating year-on-year growth beginning with the following launch of our GeForce RTX 30 Series GPUs. Based on the Ampere GPU architecture, the 30 Series has been our most successful launch ever, driving incredible demand and setting records for both desktop and laptop GPU sales. Channel inventories are still leading[ph] and we expect to remain supply constraint into the second-half of the year. With our Ampere GPU architecture now ramping across the stack in both desktop and laptops, we expect the RTX upgrade cycle to kick into high gear, as the vast majority of our GPU installed base needs to upgrade.

The laptops continue to drive strong growth this quarter, as we started ramping the Ampere GPU architecture across our line up. Earlier this month, all major PC OEMs launched 30 Series laptops based on the 3080, 3070 and 3060 as part of their spring refresh. In extreme versions based on the 3050 and 3050 Ti will be available this summer just in time for back to school, starting a price point as low as $799. This is the largest ever wave of GeForce gaming laptops over a 140 in total as OEMs address the rising demand from gamers, craters and students for Nvidia's powered laptops. The RTX biggest generational leap in performance ever. It also features our second generation (Technical Difficulty) boosting AI powered DLSS. (Technical Difficulty) For graphics with over (Technical Difficulty) 60 accelerated games. This quarter, we added many more, including Call of Duty Modern Warfare, Crysis Remastered and Outriders.

We also announced that DLSS is now available in Unreal Engine 4 and soon in the Unity game engine 4, enabling game developers to accelerate frame rates with minimal effort. The RTX 30 series also offers Nvidia Reflex, a new technology reflex is emerging as a must-have feature for Esports Gamers, who play competitor Fortnight, Valront, and Apex Legends. We estimate that about 70 (Technical Difficulty) play Esports games and (Technical Difficulty) Pros compete on GeForce. We Believe gaming also benefited from crypto mining demand, although, it's hard to determine to what extent. We've taken

actions to optimize GeForce GPUs for gamers, while separately addressing mining demand processors or CMPs.

Last week, we announced that newly manufactured GeForce RTX 3080, RTX 3070 and RTX 3060 Ti graphic cards will have their Ethereum mining capabilities reduced by half and carry a low hash rate or LHR identifier. Along with the updated RTX 3060, this should allow our partners to get more GeForce cards into the hands of gamers at better prices. To help address mining demand CMP products launched this quarter optimized for mining performance and efficiency, because they don't meet the specifications required of a GeForce CPU, they don't impact the supply (Technical Difficulty). CMP Revenue was $155 million in Q1 reported as part of the OEM and other category. And our Q2 outlook assumes CMP sales of $400 million. Our GeForce now cloud gaming platform past 10 million registered numbers this quarter.

GFN offers nearly 1,000 PC games from over 300 publishers, more than any other cloud gaming service, including 80 of the most popular free-to-game-play games. GFN expands the reach of GeForce to billions of under powered Windows PCs, Macs, iPhones and iPads. GFN is offered in over 70 countries with our latest expansions including Australia, Singapore and South America.

Moving to progress. Q1 revenue was $372 million, up 21% both sequentially and year-on-year. Strong notebook growth was driven by a new sleek and powerful RTX powered mobile workstations with Max-Q technology. And the Enterprises continue to support remote workforce initiatives. Desktop workstations rebounded as enterprise resume the spending that has been deferred during the lockdown with continued growth, likely as offices open. Key verticals driving Q1 demands include manufacturing, healthcare, automotive, and media and entertainment. At GTC, we announced the coming general availability of Nvidia Omniverse Enterprise. The world's first technology platform that enables global 3D design teams to collaborate in real-time in a shared space working across multiple software speed.

This incredible technology built on Nvidia's entire body of work and is supported by a large rapidly growing ecosystem. Early adopters include sophisticated design teams at some of the world BMW Group, Foster and Partners, and WPP. Over 400 companies have been evaluating Omniverse and nearly 17,000 users have downloaded the open data.

Omniverse is offered as a software subscription on a per user and a per server basis. As the world becomes more digital, virtual and collaborative, we see a significant revenue opportunity for Omniverse. We also announced powerful new Ampere architecture GPUs for next generation. The new RTX power work stations will be available from all major OEMs .

Moving to Automotive. Q1 revenue was a $154 million, up 6% sequentially and broken AI cockpit revenue (Technical Difficulty) it was partially offset by the expected decline in legacy entertainer. Extended our technology leadership with the announcement of the next generation Nvidia Drive Atlan SOC. Atlan will deliver an unrivalled 1,000 trillion operations per second of performance. And integrate data center class Nvidia BlueField

networking and security technologies to enhance vehicle performance and safety, making it a true data center on wheels. Atlan, which targets automakers 2025 models, will follow the Nvidia drive or an associate, which delivers 254 tops that has been selected by leading vehicle makers for production timelines starting next year.

The Nvidia drive platform has achieved global adoption across the transportation industry. Our automotive design when pipeline now exceeds $8 billion to fiscal 2027. Most recently, Volvo Cars announcement it will use, Nvidia drive orange building on our next great momentum with some of the largest automakers including Mercedes-Benz, SAIC and Hyundai Motor Group. In global taxis, we added GM cruise to the growing number of companies adopting the Nvidia drag platform, which includes Amazon Zoox and DiDi, with new energy vehicle makers on latest wins include Faraday Future or Auto, IM Motors and VinFast, which Join previously announced wins with SAIC, Neo, Extang, (inaudible) and in trucking (inaudible) is partnered with two simple in selecting Nvidia drive for autonomous driving, joining previously announced global autonomous solutions and plus. Nvidia is helping to revolutionize the transportation industry. Our full staff software divide AV and AI cockpit platform spans, silicon, systems, software and AI data center infrastructure enabling over the year upgrades to enhance safety and the joy of driving throughout the vehicles lifetime.

Starting with our lead partner, Mercedes-Benz,Nvidia drive can transform the automotive industry with amazing technologies delivered through new software and services business models. Moving to data center, revenue top $2 billion for the first time. Growing 8% sequentially and up 79% from the year-ago quarter which did not include Mellanox. Hyperscale customers lead our growth this quarter as they built infrastructure to commercialize AI in their services. In addition, cloud providers have adopted the a 100 to support growing demand for AI from enterprises, startups and research organizations. Customers have deployed Nvidia's a 100 DGX platforms to train deep neural networks with rising computational intensity led by two of the fastest growing areas of AI, natural language understanding and get deep recommendaters with deep recommendaters.

In March, Google cloud platform announced general availability of the A100 with early customers, including square for its cash application and alphabets deep mind. The A100 is deployed across all major hyperscale and cloud service providers globally. And we see strengthening demand in the coming quarters. Every industry is becoming a technology industry, an accelerating investments in AI infrastructure both through the cloud and on-premise.

Our vertical industries grew both sequentially and year-on-year, led by consumer internet companies. For example may there a leading internet technology company in Korea and Japan, this training giant AI language models outscale on DGX SuperPOD to pioneer new services across e-commerce, search, entertainment, and payment applications. We continue to gain traction in inference with Hypersale and vertical industry customers across a broadening portfolio of GPUs. We had record shipments of GPUs in inference growth is driving, not just the T4, which runs up strongly in the quarter, but also the universal A100 core GPU as well as new both the new and peer architecture based A10 and A30 GPUs. All excellent at training as well as inferencing.

Customers are increasingly migrating from CPUs to GPUs for AI inferencing for two cheap reasons: First, GPUs can better keep up with the exponential growth in the size and the complexity of deep neural networks and respond with the required low latency. In April, -- and our first AI inferencing benchmark Nvidia achieved the top results across every category. Spending computer vision, medical imaging, recommend data systems, speech recognition and natural language processing.

And second, Nvidia's full-stack inference platform including triton inference server software, simplifies the complexity of deployment AI applications by supporting models from all major frameworks and optimizing for different query types, including batch real-time and streaming. Triton is supported by several partners in the cloud services including Amazon, Google, Microsoft and Tencent. Examples of how customers use Nvidia's inference platform include Microsoft for grammar checking and office, the United States Postal Service to real-time package analytics, T-Mobile for customer service, Pinterest board, image search, and GE Healthcare for heart disease detection.

We also have strong results with non auction networking products, like our compute business, strong growth was driven by hyperscale customers across both Ethernet and InfiniBand, breaches key design wins and proof-of-concept trials for the Nvidia BlueField-2 DPU, the cloud service providers and consumer internet companies. We also unveiled BlueField-3, the first DPU built for AI and accelerated computing, with support from VMware, NetApp, Splunk, Cloudflare and others.

BlueField-3 is the industry's first 400 Gigg DPU delivers the equivalent data center services of up to 300 CPU cores. It transforms traditional server infrastructure into zero trust environment, in which every user is authenticated by offloading and isolating data center services for business applications. With BlueField-3 our DPU roadmap will deliver an unrivalled 100x performance increase over a three-year period.

As we look back at the first full-year since closing the Mellanox acquisition, we are extremely pleased with how the business has performed. It has not only exceeded our financial projections, but it has been instrumental in key new platforms like the DGX SuperPOD and the BlueField DPU, enabling our data center scale computing strategy.

In April, we held our largest ever GPU technology conference and with more than 200,000 registrants from 195 countries, Jensen's came out had over 14 million views. At GTC, we announced our first data center CPU, Nvidia (inaudible) targeted at processing massive next generation AI models with trillion to parameters the arm-based processors processor will enable 10x the performance and energy efficiency of today's faster servers. With Grace Nvidia has a three-chip strategy with GPU, DPU and now CPU. A Swiss National Supercomputing Center and the U.S. department of energy was almost national laboratory are the first to announce plans to build grace powered supercomputers. Grace will be available in early 2023.

GTC is first and foremost for developers. We announced Nvidia developed and optimize pre-trained model availability on the Nvidia GPU cloud registry. Developers can choose a pre-trade model and adapted to fit their specific needs using Nvidia TAO, our transfer

learning software. TAO fine tunes the model with customers own small data set to get models a custom fit without the cost, time and massive data sets required to train a neural network from scratch. Once a model is optimized and ready for deployment, users can integrate it with the Nvidia application framework that fits their use.

For example, the Nvidia Jarvis framework for interactive conversational AI is now generally available and used by customers such as T-Mobile and Staff. And the Nvidia merlin framework for deep recommendator with an open beta with customer such as now in Tencent. With the chosen application framework users can launch Nvidia fleet command software to deploy and manage the AI application across a variety of Nvidia GPU powered devices.

For Enterprise customers, we unveiled a new enterprise grade software offerings available as a perpentual license or subscription NVIDIA AI enterprise is a comprehensive suite of AI software, that speeds development and deployment of AI workloads, and simplifies management of enterprise AI infrastructure through our partnership with VMware hundreds of thousands of customers will be able to purchase Nvidia AI Enterprise, with the same from earlier pricing model that IT managers used to procure the VMware infrastructure software. We also made several announcements awaiting the delivery of both Nvidia AI and accelerated computing to enterprises and end users among the world's largest industries, leading several OEMs launched Nvidia certified systems which are industry standard servers based on the Nvidia EGX platform, they run Nvidia AI enterprise software center and are supported by the Nvidia A30 and A10 GPUs. Initial costumers including (inaudible) and last general brigham.

In addition, we announced the Nvidia AI on 5G platform supported on Nvidia EGX service to 5G RAM and AI applications. The AI on 5G platform leverages the Nvidia aerial software and the Nvidia BlueField too 100 converge car, which combined our GPUs and DPUs. We are teaming with Fujitsu, Google cloud,Mavenir, Radisys and Wind River into AI on 5G platform to speed the creation of smart cities and factories, advanced hospitals, and intelligence stores.

Another highlight, at GTC was the announcement of a broad range of initiatives to strengthen arm ecosystem across cloud data centers, HPC, Enterprise and Edge and PCs.

In the cloud, we are bringing together AWS Graviton2 processors and Nvidia GPUs to provide a range of benefits, including lower costs support form return gains extremely experiences and greater performance for arm-based workloads.

In HPC, we are bringing together an Ampere wth Nvidia GPUs and Nvidia HPC software development kit. Special-- super supercomputing centers deploying it, include Oak Ridge and Los Alamos National Labs. In this sent -- in the Enterprise and Edge, we are bringing together Marvel arm-based OCTEON processors and the Nvidia GPUs to accelerate video analytics, and cybersecurity PCs we are bringing together media text arm-based processors within videos RTX GPUs s to enable realistic, raytrace graphics and cutting-edge AI laptop.

FINAL

Bloomberg Transcript

On our arm at this acquisitions. We are making steady progress in working with the regulators across key regions. We remain on track to close the transaction within our original time frame of early 2022, arms IP is wide reused but the company needs a partner that can help it achieve new heights. Nvidia is uniquely positioned -- and we are committed to invest in developing the arm ecosystem, enhancing R&D, adding IP and turbo charging is development to grow into new markets to and embedded devices areas, where only has a light footprint, or in some cases not at all.

Moving to the rest of the P&L. GAAP gross margin for the first quarter was down 100 basis points from a year earlier and up 100 basis points sequentially. Non-GAAP gross margin was up 40 basis points from a year earlier and up 70 basis points sequentially. The sequential non-GAAP increase was largely driven by a more favorable mix within data center and the addition of CMP products. Q1 GAAP $3.03, up 106% from a year earlier, non-GAAP EPS was $3.66, up 103% from a year-ago. Q1 cash flow from operations was $1.9 billion.

Let me turn to the outlook for the second quarter of fiscal 2022. We expect broad-based sequential year-on-year revenue growth in all of our market platforms. Our outlook include $400 million in CMP. Aside from CMP, the sequential revenue increase in our Q2 outlook is driven largely by data center and gaming. In data center we expect sequential growth in both compute and networking. In gaming with the move to low hash rate GeForce GPUs and increasing the amount of CMP products, we are making a significant effort to serve minors with CMPs and provide more (inaudible) to Gamers. If there is additional CMP demand we have supplied flexibility to support it. We believe these actions combined with strong gaming demand will drive an increase in our core gaming business for Q2.

Now to look at our outlook for Q2. Revenues expected to be $6.3 billion, plus or minus 2%. GAAP and non-GAAP gross margins are expected to be 64.6% and 66.5% respectively, plus or minus 50 basis. GAAP and non-GAAP operating expenses are expected to be approximately $1.76 billion and $1.26 billion respectively. GAAP and non-GAAP, other income and expenses are both expected to be an expense of approximately $50 million. GAAP and non-GAAP tax races are both expected to be 10%, plus or minus 1%, excluding discrete items.

Capital expenditures are expected to be approximately $300 million to $325 million. Further financial details are included in the CFO commentary and other information available on our IR website.

In closing. Let me highlight (inaudible) and then of EBITDA[ph] will keynote (inaudible) on the evening of May 31 U.S. time as well as several upcoming events for the financial community will be virtually attending the EverCore TMT conference on June 7th, the base 2020 global technology conference on June 9th and the NASDAQ Virtual Investor Conference on June 16th. Our earnings calls to discuss our second quarter results is scheduled for Wednesday August 18th.

With now, we will open the call for question. Operator, would you please pole for questions?

Questions And Answers

Operator

(Question And Answer)

(Operator Instructions) And your first question comes from to Timothy Arcuri with UBS.

Q - Timothy Arcuri {BIO 3824613 <GO>}

Thanks a lot. Colette, I was I was wondering if you can double-click a little more on the guidance? I know of the 600 to 650 in growth, you said 250 is coming from CMP and both gaming and data center will be up. I -- can we assume that they're up about equally so you're getting about 200 roughly from each of those? And I guess second part of that is, within data center, I am wondering can you speak to the networking piece? It sounds like maybe it was up a bit more modestly than it's been up the past few quarters? I am just wondering what the outlook is there. Thanks.

A - Colette Kress {BIO 18297352 <GO>}

Yes. Thanks so much for the question on our guidance. So I just want to start off with we see demand really across all of our markets, all of our different market pipelines, we do plan to grow sequentially. You are correct that we are expecting increase in our CMP, and outside of our CMP growth, we expect a lion share of our growth to come from our data center and gaming. In our data center business, right now, our product line up couldn't be better. We have a strong overall portfolio both for training and for in inferencing and we are seeing strong demand across our hyperscales and vertical industries.

We've made a deliberate effort on the gaming perspective to supply to our gamers the cards that they would like given the strong demand that we see. So that will also support the sequential growth that we're receiving. So you are correct that we do see a growth sequentially coming from datacenter and gaming both contributing quit well to our growth.

Q - Timothy Arcuri {BIO 3824613 <GO>}

Thanks a lot, Colette.

A - Colette Kress {BIO 18297352 <GO>}

I didn't answer your second question, my apologies on Mellanox. Additionally, Mellanox is an important part of our data center, it is quite integrated with our overall products. We did continue to see growth this last quarter and we are also expecting them to sequentially grow as we move into Q2. They are smaller part of our overall data center business, but again, we do expect them to grow.

Operator

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

Q - C.J. Muse

Yes, good afternoon, thank you for taking the question. In your prepared remarks, I think I heard you talk about a vision for acceleration in data center as we go through the year. And as you think about the purchase obligations that you reported up 45% year-over-year, how much of that is related to long lead time data center and how do should we interpret that in terms of what kind of ramp we could see in the second half, particularly as you think about perhaps adding more growth from enterprise on top of what was hyperscale driven growth in the April quarter? Thank you.

A - Colette Kress {BIO 18297352 <GO>}

So let me -- regarding our purchasing, our purchasing of inventory, and what we're seeing in just both our purchase commitment general inventory. The market has definitely changed to where long lead times are required to build out our data center products. So we're on a steady stream longer term, so that we can make sure that we can serve our products that we have. So yes, a good part of those purchase commitments is really about the long lead times of the components to create the full systems. I'll turn the second part of the question over to Jensen.

A - Jensen Huang {BIO 1782546 <GO>}

What was the second part of the question, Colette?

A - Colette Kress {BIO 18297352 <GO>}

Second part of the question was what do we see in the second half as it relates to the lineup of Enterprise we articulated in our pre remarks regarding that we're seeing an acceleration. Thank you.

A - Jensen Huang {BIO 1782546 <GO>}

Yes. We're seeing the strength across the board in data centers and we're seeing strengthening demand. Our data center as you know is always range of applications from scientific computing, both physical and life sciences, data analytics, and classical machine learning. Cloud computing and cloud graphics, which is becoming more important because of remote work. And very importantly, AI, both for training as well as inferencing for classical machine learning models like XGBoost all the way to deep learning based models like conversational AI, natural language understanding, recommender systems and so on.

And so we have a large suite of applications, and our NVIDIA AI and NVIDIA HPC SDKs accelerate these applications and data centers. They run on systems that range from HGX for the hyperscalers to DGX for on-prem to EGX for enterprise and edge, all the way out to AGX autonomous systems. And this quarter at GTC, we announced of our largest initiatives and it's taken us several years. You've seen working on it in open -- on the open

over the course of the last several years, and it's called EGX, it's our enterprise AI platform. We're democratizing AI, we're bringing it out in cloud, we're bringing it to enterprises and we're bringing it to the edge. And reason for that is because the vast majority of the world's automation that has to be done has data that has data sovereignty issues or data raid issues that can't move to the cloud easily.

And so we got -- we have to move the computing to the premise and oftentimes all the way out to the edge. The platform has to be secure, has to be confidential, it has to be remotely manageable, and of course that has to be high performance and has to be cognitive and has to be built like the cloud, the modern way of building cloud data centers.

And so these stacks has to be monitoring on the one hand and has to be integrated into classical enterprise systems on the other hand, which is the reason why we work so closely with VMware and accelerated VMware's operating systems -- data center operating systems, software defined datacenter stacks on BlueField. Meanwhile, we ported in NVIDIA AI, NVIDIA HPC on to VMware so that they could run, distribute a large scale accelerating computing for the very first time. And that partnership was announced at VMworld, it was announced at GTC, and we're in the process of going to market with all of our enterprise partners, the OEMs, the value-added resellers, their service -- their solution integrators all over the world.

So this is a really large endeavor and the early indications of it are a really exciting. And the reason for that is because as you know, our data center business is 50% vertical industry enterprises already. It's more than 50% vertical industry enterprises already, and by creating this easy-to-adapt and easy-to-integrate stack, it's going to allow them to move a lot a lot faster. And so this is the next major wave of AI, this is a very exciting part of our initiative, and the something I've been working on for we've been working on for quite a long time. And so I'm delighted with the launch this quarter at GTC.

The rest of the data center is doing great too. As Colette mentioned, hyperscale demand is strengthening, we're seeing that for computing and networking. And you know that the world's cloud data centers are moving to people only, because every small percentage that they get out of predictive inference, drives billions and billions of dollars of economics for them. And so the movement towards deep learning shifts the data center workload away from CPUs, because accelerators are so important. And so hyperscale, we're seeing great traction and great demand. And then lastly, supercomputing. Supercomputing centers all over the world are building out, and we're really in a great position there to fuse for the very first time simulation-based approaches as well as data driven based approaches what is called artificial intelligence.

And so across the board, our data center is gaining momentum. And we see -- we just see great strength right now and it's growing strength, and what's really set up for years of growth in data center. This is the largest segment of computing as you know. And this segment of computing is going to continue to grow for some time to come.

Operator

And your next question comes from Aaron Rakers with Wells Fargo.

Q - Aaron Rakers {BIO 6649630 <GO>}

Yes, thanks for taking the question. Congratulation on the results. I'm going to try slip in two of them here. First of all, Colette, I think in the past you talked about how much of your gaming install base is kind of on the pre ray tracing platforms are really kind of in context behind the upgrade cycle, that's still in front of us? That's kind of question one. And then on the heels of the question, I'm just curious things like VMware's Project Monterey as we think about the BlueField-2 product and BlueField-3, how should we think about those starting to become or when should they become really material incremental revenue growth contributors for the company? Thank you.

A - Colette Kress {BIO 18297352 <GO>}

So yes, we have definitely discussed in terms of the great opportunity that we have in front of us, of folks moving to our ray-traced GPUs and we're in the early stages of that. We've had a strong cycle already, but still, we probably have approximately 15% moving up a little bit from that at this time. So it's a great opportunity for us to continue to upgrade a good part of our install base. Not only just with our desktop GPUs, but the RTX laptops are also a great driver of growth and upgrading folks to our RTX.

A - Jensen Huang {BIO 1782546 <GO>}

Colette, do you want me to take the second one?

A - Colette Kress {BIO 18297352 <GO>}

Yes, please.

A - Jensen Huang {BIO 1782546 <GO>}

Aaron, a great question on BlueField. First of all, the modern data center has to be re-architected for several reasons. There are several fundamental reasons that makes it very, very clear that the architecture has to change. The first -- so it's cognitive, which means that a data center is shared for everybody, it's not dependent, you don't know who's coming and going and it's exposed to everybody on the internet.

Number two, you have to assume that it's a zero trust environment, right? Because you don't know who's using it. It used to be that we have perimeter security, but those days are gone. Because it's cognitive, its remote access, its multi-tenant, your -- it's public cloud, the infrastructure is used for internal and external applications. So number two has to be -- it has to be zero trust.

The third reason is something that started a long time ago, which is software defined in every way, because you don't want a whole bunch of bespoke custom gear inside a data center, you want to operate the data center, what software you want to be software-defined? The software-defined data center movement enabled this one pane glass a few IT managers orchestrating millions and millions of computers at one place. And the software that -- the software runs what used to be storage, networking, security, virtualization, and all of that -- all of those things have become a lot larger and a lot more

intensive. And it's consuming a lot of the data center. In fact, you estimate depending on how you want to think about it, how much security you want to want to put on it. If you assume that is a zero trust data center, probably half of the CPU cores inside that data center is run not applications.

I mean that's kind of strange, because you created the data center to run services and applications, which is the only thing that makes money. The other half of the computing is completely soaked up running the software-defined data center just to provide for those applications. And that you could imagine even accepting if you like as the cost of doing business, however, it co mingles the infrastructure, the security point and the application point, and then exposes the data center to attackers. And so you fundamentally want to change the architecture as a result of that to offload that software-defined virtualization and the infrastructure operating system if you will and the security services to accelerate it, because Moore's law has ended and moving software, that was running on one set of CPUs which is really, really good already to another set of CPUs that were going to make it more effective. Separating it doesn't make it more effective. So if you want to offload that and take the -- take that application and software and accelerate it using accelerators, a form of accelerated computing.

And so that's -- these things are fundamentally what BlueField is all about and we created the processor that allows us to -- BlueField-2 replaces approximately 30 CPU cores, BlueField-3 replaces approximately 300 CPU cores to just -- to put it, give you a sense of it. And then BlueField-4, we're in the process of building already. And so we've got a really aggressive pipeline to do this. Now, how big is this market? So the way you're thinking about that is every single networking chip in the world will be a smart networking chip. It will be a programmable, accelerated, infrastructure, processor. And that's what the GPU is, is a data-center on a chip.

And I believe every single server node will have it, it will replace today's mix with something like BlueField, and it will offload about half of the software processing that's consuming data centers today. But most importantly, it will enable this future world where every single packet, every single application is being monitored in real-time all the time for intrusion. And so how big is that application? How big is the market? Just 25 million servers a year, that's the size of the market. And we know that servers are growing and so just to give you a feeling for that.

And then the future servers are going to move out to the edge, and all of those edge devices will have something like BlueField. And then how are we doing? We're doing POCs now with just about every internet company. We're doing really exciting work there. We've included it in high-performance computing, so that it's possible for supercomputers in the future to be cognitive, to be zero trust, to be secure, and still be a super computer. And then we expect next year to have a meaningful, if not significant, revenues contribution from BlueField, and this is going to be a really large growth market for us. You can tell, I'm excited about this and I've put a lot of energy into it, the company is working really hard on it. And this is a form of accelerated computing that's going to really make a difference.

Operator

And your next question comes from Vivek Arya with Bank of America Securities.

Q - Vivek Arya {BIO 6781604 <GO>}

Thanks for taking my question. Jensen, is NVIDIA able to ring fence this crypto impact in your CMP product. So even if let's say crypto goes away for whatever reason, the decline is a lot more predictable and manageable than what we saw in the 2018-'19 cycle. And then kind of Part B of that is how do you think about your core PC Gamer demand? Because when we see these kind of 106% year-on-year growth rate, it brings questions of our sustainability. So give us your perspectives on these two topics. Just how does one ring fence kind of the crypto effect and what do you think about the sustainability of your core PC gamer demand? Thank you.

A - Jensen Huang {BIO 1782546 <GO>}

Sure. Thanks, Vivek. First of all, it's hard to estimate exactly how much and where crypto mining is being done. However, we can only assume that the vast majority of it is contributed by professional miners, especially when the amount of mining increases tremendously like it has. And so we created CMP, and CMP and GeForce are not fungible. You could use GeForce for mining, but you can't use CMP for gaming. CMP is yields better, and produces those, that will take away from the supply to GeForce. And so it protects our GeForce supply for the gamers. And the question that you have is what happens when on the tail end of this?

There are several things that we hope and we learned a lot from the last time, but you never learn enough about this dynamic. What we hope is that the CMPs will satisfy the miners and will stay in mines, in the professional mines, and we're trying to produce a fair amount of them, and we have, we secured a lot of a lot of demand for the CMPs and we'll fulfill it. And what makes it different this time is several things. One, we're in the beginning of our RTX cycle whereas Pascal was the last GTX. And not only that, it was at the tail end of the GTX cycle. It was the last GTX and it was a tail end of GTX cycle.

And we're at the very beginning of the RTX30 cycle. And because we reinvented computer graphics, we reset the computer industry. And after three years, the entire graphics industry has followed. Every game developer has moved to ray tracing, every content developer and every content tool has moved to ray tracing. And so if you move to ray tracing, these applications are so much better and they simply are too slow on GTXs. And so we're seeing a reset of the install base if you will. And at a time when the gaming market is the largest ever, we've got this incredible install base of GeForce users. We reinvented computer graphics and we've reset their -- reset the install base and created an upgrade opportunity that's really exciting, at a time when the market is -- the gaming market, the gaming industry is really large. And what's really exciting on top of that is that gaming is no longer just gaming, but it's infused into sports, e-sports, it's infused into art, it's infused into social. And so gaming has such a large cultural impact now, it's the largest form of entertainment. And I think that the experience we're going through is going to last a lot. And so one I hope that crypto will -- the CMP will steer our GeForce supply to gamers where we see strong demand and I expect to see a strong demand for quite some time, because of the dynamics that I described. And hopefully, in the combination with this two, we'll see a strong growth and through strong growth in our core gaming business through the year.

Operator

And your next question comes from John Pitzer with Credit Suisse.

Q - John Pitzer {BIO 1541792 <GO>}

Yes. Good afternoon, guys. Thanks for letting me ask the question. And I have two hopefully quick questions. First, I hearken back to the mantra you guys put out a couple of analyst days ago, the more you spend the more you save, and you've always been very successful as you brought down the cost of doing something to really drive penetration growth. And so I'm curious with the NVIDIA Enterprise AI software stack, is there a sense that you can give us as how much that brings down the cost of deployment in AI inside the Enterprise. And do you think whether COVID lockdown related or cost related -- there's pent-up demand that this unlocks?

And then my second question is just around government subsidies. A lot of talks out of Washington about subsidizing the chip industry, a lot of that goes towards building fabs domestically. But when I look at AI, I can't think of anything more important to maintain sort of leadership in relative to national security. How do we think about Nvidia and kind of the impact that these government subsidies might have on either you or your customers or your business trends?

A - Jensen Huang {BIO 1782546 <GO>}

The more you buy, the more you shall save, there's no question about that. And the reason for that is because we're in the business of accelerated computing. We don't accelerate every application, however, for the applications we do accelerate, the acceleration is so dramatic, and because we sell a component, the entire system, the TCO, the entire system and all the service and all the people and the infrastructure and the energy costs has been reduced by X factors, sometimes 10x, sometimes 15x, sometimes 5x. And so the -- when we set our mind on accelerating a certain class of applications, and recently, we worked on cuQuantum so that we could help the quantum industry, quantum computing industry accelerators, simulators so that they could discover new algorithms and invent future computers. And even though it won't happen until 2030 for the next 20 years, we're going on 15 years -- we're going to have some really, really great work that we can do using Nvidia GPUs to do quantum simulations.

We recently did a lot of work in natural language understanding in computational biology so that we could decode biology and understand how biology is to infer to understand it, and to predictively improve upon it, and design new proteins, those words are so vital and that's what accelerated computing is all about.

Our enterprise software -- and I really appreciate the question -- our enterprise software used to be just about the VGPU, which is virtualizing GPU inside the VMware environment or inside the RedHat environment and makes it possible for multiple users to use one GPU which is the nature of enterprise virtualization, but now with Nvidia AI, Nvidia Omniverse, Nvidia Fleet Command, whether you are doing collaboration or virtual simulations for robotics and digital twins, designing your factory or you're doing data analytics learning what the predictive features are that could create an AI model, predictive model that you

can deploy out at the edge using fleet command. We now have an end-to-end suite of software that is consistent with today's enterprise service agreements.

It's consistent with today's enterprise business models and allows us to support customers directly and provide them with the necessary service promises that they expect, because their delivery -- they're trying to build a mission critical application on top. And more importantly, by creating this product timing or software, we provide the ability for our large network of partners, OEM partners, value-added resellers, system integrators, solution providers for this large network of hundreds of thousands of IT sales professionals that we are connected to through our network, we give them a product that they can take the market.

And so the distribution channel, the sales channel of VMware, the sales channel, all of Cloudera, the sales channel of all of our partners and EDA and design, Autodesk (inaudible) so on and so forth, all of these sales channels and all of these partners are now partners in taking our stacks to the market. And we have a fully-integrated system that are open to the OEMs so that they could create systems that run the stack. And so all certified, all tested, all benchmark, and of course, very importantly, all supportive. And so this new way of taking our products to market, whereas our cloud business is going to continue to grow and that's -- that part of AI is going to continue to grow, and that business is direct. We sell components directly to them, we support them directly, but there are 10 of those customers in the world. For enterprises, there are thousands, industries far and wide. And so I think this -- we now have a great stack and a great software stack that kind of allows us to take it to the world's market, so that everybody could buy more and save more.

Operator

And our final question comes from Stacy Rasgon with Bernstein.

Q - Stacy Rasgon {BIO 16423886 <GO>}

Hey, guys, thanks for taking my questions. I think it's for Colette. So Colette, last quarter, you had kind of suggested that Q1 would be the task for I guess for gaming as well as the rest of the company, the gaming in particular, and it would grow sequentially through the year. I guess given the strength we're seeing in the first half, do you still believe that, that is the case. And I kind of heard you guys, I think, kind of dance around that points a little bit to your response to one of the other questions, but could you clarify that, is that still your belief that core gaming business can grow sequentially through the rest of the year.

And I guess same question is for center data center, especially since sounds like hyperscale is now coming back like after a few quarters of digestion and then all of the other tailwinds you've talked about. I mean is there any reason to think the data center itself shouldn't also grow sequentially through the rest of the year?

A - Colette Kress {BIO 18297352 <GO>}

Yes, Stacy, thanks for the question. So I first want to start with -- when we talked about our Q1 results, and when we were looking at Q1, we were really discussing a lot about what we expected between Q4 and Q1. Given what we knew was still high demand for gaming, we

believe we will continue to grow between Q4 and Q1, which often we don't. And we absolutely have the strength of overall demand to grow.

What that then led was again continued growth from Q1 to Q2 as we are working hard to provide more supply for the strong demand that we see. We have talked about that we have additional supply coming. We expect to continue to grow as we move into the second half of the year as well for gaming. Now we only guide one quarter at a time, but our plan is to take the supply, serve the overall gamers work on building up the channel as we know the channel is quite lean. And so yes, we do and still expect growth in the second half of the year particularly when we see the lineup of games and the holiday overall coming back to school, all the very important cycles for us, and there's a great opportunity to upgrade this RTX install base.

Now, in terms of data center, we'll look in terms of our guidance here. We have growth from Q1 to Q2 planned in our overall guidance. And we do see as things continue to open up a time to accelerate in the second half of the year for data center. We have again a great lineup of products here, couldn't be a better line up now that we've also added the inferencing products and the host of overall applications that are using our software that we have. So this could be an opportunity as well to see that continued growth. We will work in terms of serving the supply that we need for both of these markets. But yes, we can see definitely growth in the second half of the year.

Operator

There are no further questions at this time. CEO, Jensen Huang. I'll turn the call back over to you.

A - Jensen Huang {BIO 1782546 <GO>}

Well, thank you. Thank you for joining us today. Nvidia Computing platform is accelerating. Launched at GTC, we are now ramping new platforms and initiatives. There are several that I'll mention. First, enabled by the fusion of NVIDIA RTX, NVIDIA AI, NVIDIA PhysX, we built Omniverse, a platform for virtual collaboration and virtual worlds to enable tens of millions of artists and designers to create together in their own metaverses.

Second, we laid the foundation to be a three-chip data center scale computing company with GPUs, DPUs and CPUs. Third, AI is the most powerful technology force of our time. We partner with cloud and consumer internet companies to scale out and commercialize AI-powered services, and we're democratizing AI for every enterprise in every industry. With an NVIDIA EGI certified systems, the NVIDIA Enterprise AI suit pre-trained models for conversational AI, language understanding, recommender systems, and our broad partnerships across the IT industry, we are removing the barriers for every enterprise to access state-of-the-art AI.

Fourth, the work of a NVIDIA CLARA in using AI to revolutionize genomics and biology is deeply impactful for the healthcare industry, and I look forward to telling you a lot more about this in future. And fifth, the electric, self-driving and software-defined car is coming. With NVIDIA Drive, we are partnering with the global transportation industry to reinvent the car architecture, reinvent mobility, reinvent driving and reinvent the business model of

the industry. Transportation is going to be one of the world's largest technology industries.

From gaming, metaverses, cloud computing, AI, robotics, self-driving cars, genomics, computational biology, Nvidia is doing important work and innovating in the fastest-growing markets today. As you can see, on top of our computing platforms that span PC, HPC, cloud, enterprise to autonomous edge, we've also transformed our business model beyond chips, NVIDIA vGPU, NVIDIA AI Enterprise, NVIDIA Fleet Command and NVIDIA Omniverse as enterprise software license and subscription to our business model. And NVIDIA GeForce now and NVIDIA Drive with Mercedes Benz as the lead partner, our end-to-end services on top of that.

I want to thank all of the NVIDIA employees and partners for the amazing work they're doing. We look forward to updating you on our progress next quarter. Thank you.

Operator

This concludes today's conference call. You may now disconnect.

FINAL

Bloomberg Transcript