# Q4 2021 Earnings Call

## Company Participants

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

## Other Participants

- Aaron Rakers, Analyst
- C.J. Muse, Analyst
- John Pitzer, Analyst
- Mark Lipacis, Analyst
- Stacy Rasgon, Analyst
- Timothy Arcuri, Analyst
- Vivek Arya, Analyst

## Presentation

### Operator

Good afternoon, my name is [ph]Maryama and I will be your conference operator today. At this time, all lines have been placed (technical difficulty). After the speakers' remarks, (technical difficulty). (Operator Instructions) Thank you. I will now turn the call over to Simona Jankowski, NVIDIA, Vice President of Investor Relations and Strategic Finance to begin the conference.

### Simona Jankowski  {BIO 7131672 <GO>}

Thank you. Good afternoon everyone and welcome to NVIDIA's conference call for the fourth quarter of fiscal 2021. With me on the call today from NVIDIA are Jensen Huang, President and Chief Executive Officer; and Colette Kress, Executive Vice President and Chief Financial Officer. I'd like to remind you that our call is being webcast live on NVIDIA's investor relations website. The webcast will be available for replay until the conference call to discuss our financial results for the first quarter of fiscal 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. For a discussion of factors that could affect our future financial results and business, please refer to the disclosure in today's earnings release, our most

recent Forms 10-K and 10-Q and the reports that we may file on Form 8-K with the Securities and Exchange Commission.

All our statements are made as of today, February 24, 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, Q4 was another record quarter, with revenue exceeding $5 billion and year-on-year growth accelerating to 61%. Full year revenue was also a record at $16.7 billion, up 53%. Our gaming business have reached record revenue of $2.5 billion in Q4, up 10% sequentially and up 67% from a year earlier. Full year gaming revenue was a record at $7.8 billion, up 41%.

Demand is incredible for our New GeForce RTX 30 Series products based on the NVIDIA ampere GPU architecture. In early December, we launched the GeForce RTX 3060TI which joined the previously launched RTX 3090, 3080 and 3070. The entire 30 series lineup has been hard to keep in stock and we exited Q4 with channel inventories even lower than when we started. Although we are increasing supply, channel inventories will likely remain low throughout Q1.

GeForce RTX 30 Series graphics cards were a holiday sensation. Due not just to their amazing performance, but also to their rich features, including our second-generation RTX Ray Tracing technology and DLSS AI powered performance accelerator, which massively boosts frame breaks in graphically demanding titles. 3,000 games now support RTX, including the top battle royale game, Fortnite, the top role-playing game, Cyberpunk 2077, the top massively multiplayer online game, World of Warcraft, and the best selling game of all time, Minecraft. RTX has clearly set the new standard in gaming.

Building on this momentum, at CES in January we introduced a wave of ampere architecture gaming products, including our biggest ever top launch powered by GeForce RTX 3060, 3070 and 3080 laptop GPUs and with our third generation Max-Q technology. These new thin and lightweight gaming laptops increased performance and energy efficiency by up to 2x from the prior generation. RTX 3060 laptops start $999 and are faster than the previous generation laptops which sold for $2,500. The incredible performance, design and price points of these new laptops will delight the growing universe of gamers and creators as well as students and professionals.

The gaming laptop market has grown seven fold in the past seven years and momentum is building, with top OEMs bringing to market a record 70 plus laptop models based on the GeForce RTX 30 Series. GeForce laptops as a whole, are the fastest growing and one of the largest gaming platforms.

Onto at CES, we announced the GeForce RTX 3060 GPU, priced at $329, extending the 30 Series desktop (inaudible) further into the mainstream. We expect strong demand when it launches this Friday as 60 class GPUs have traditionally been our most popular products. Starting with the 3060, we're taking an important step to maximize the supply of GeForce GPUs for gamer. Users are constantly discovering new applications for our powerful programmable GPUs and cryptocurrency mining is one of them

With rising Ethereum prices, there are indications that miners are behind GPUs. We would like GeForce GPUs to end up with gamers. So we have created a new special software drivers that will detect the Ethereum lining algorithm, cutting in half the mining efficiency of the GeForce RTX 3060. We suspect the significant increase in the Ethereum network hash rate observed over the past few months was driven by a combination of previously installed mining capacity that was reactivated as well as new sales of GPUs and ASICs. Since our GPUs are sold to graphics card manufacturers and then on to distribution, we don't have the ability to accurately trap or quantify their end use. And those estimates suggest that crypto mining contributed $100 million to $300 million to our Q4 revenue, or a relatively small portion of our gaming revenue in Q4.

Crypto currencies have recently started to be accepted by companies and financial institutions and show increased signs of staying power. To address industrial Ethereum mining demand, last week, we announced a new line of NVIDIA CMPs or crypto mining processors. Shipments will start in March. CMPs lack display outputs and have other optimizations that improve crypto mining power efficiency. CMP products will let us gain some visibility into the contribution of crypto mining to our overall revenue. For Q1, we estimate that CMP will contribute approximately $50 million. We plan to sell these products to industrial miners. We will quantify their contribution each quarter for transparency.

Over the past year, it has become clear that we've entered a new era in which gaming is an integral part of global culture. The number of concurrent users unseen has more than doubled since 2018 and continues to hit new records. In 2020 alone, more than 100 billion hours of gaming content was seen on YouTube, and half a billion people watch eSports. Increasingly, we aren't just gaming. We're also watching sports, attending concerts, creating content and connecting with our friends in virtual environments.

Additionally, we are excited about the new experiences like (inaudible). Significantly more content is now available, including arguably the first VR killer app, (inaudible), and there is now almost 2 million VR users unseen. And with these powerful structural shifts, we expect our gaming business to remain on a robust growth trajectory. The GeForce RTX 30 Series GPUs have kicked off a powerful upgrade cycle and we estimate only around 15% of GeForce gamers own an RTX (inaudible) GPU, which is needed to experience the beautiful ray traced graphics of modern games. Moreover, the universe of gamers is rapidly expanding and the reach of GeForce has extended beyond gamers to some 45 million creators.

In addition, gaming revenue continues to benefit from a favorable mix shift as gamers and creators keep moving to higher end GPUs, we expect another great year for GeForce. Earlier this month, we celebrated the one-year anniversary of the GeForce NOW cloud

gaming platform, which is now over 6 million members strong. GeForce NOW offers 800 PCs from over 300 publishers, more than any other cloud gaming, service including 80 of the most played free to play games. Starting with support for our Windows PCs, Macs and (inaudible) devices, we added support in recent months to Chromebooks iPhones and iPad's. GFN has grown globally with more than 65 countries on our service and more added regularly by our GeForce NOW alliance partners.

Moving to ProViz, Q4 revenue was at $307 (technical difficulty) down 10% year-on-year and ahead of our expectations. Full-year revenue was $1.1 billion, was down 13%. Strong sequential growth was driven primarily by a recovery in desktop workstations as some customers returned to the office and enterprises resumed purchases that had been deferred by the pandemic. Notebook GPUs grew sequentially to a record as enterprises continued to support remote workforce initiatives.

Looking ahead, the reopening of businesses will benefit desktop workstations. The longer-term workforce trends will likely shift of next to notebook GPUs and cloud offerings. Healthcare was stand out vertical in the quarter, with significant orders from GE, Semens, and Oxford Nanopore Technologies, public sector and automotive also showed strength.

Omniverse, our real-time 3D collaboration and simulation platform is now an open beta. Over 500 creators and professionals have tested Omniverse through our early access program. Omniverse is one of our most important and exciting platforms. We are delighted by its initial acceptance and look forward to sharing more details on its long-term growth opportunity in the coming months.

Moving to Automotive. Q4 revenue was $145 million, up 16% sequentially and down 11% year-on-year. Full year revenue of $536 million declined 23%. Sequential growth was driven by continued recovery in the global automotive production volumes and growth in AI cockpit revenue. The year-on-year decline reflects the expected ramp down of legacy infotainment. NVIDIA has emerged as the industry's leading end-to-end full stack technology provider for self-driving and AI enabled vehicles. Orin the SOC that drive self driving platform was built on, delivers an unrivaled 254 trillions of operations per second. A performance on industry leading power efficiency, helping to revolutionize the transportation industry.

Our technology leadership has driven a robust rapidly growing set of opportunities. We have great momentum, with an expanding list of electric vehicle OEMs including [ph]VO, SAIC, [ph]Viado and [ph]X-Teng[ph] which are all using the NVIDIA DRIVE platform to power their next generation of vehicles. We look forward to growing with them as they continue to scale.

Our software defined platforms is the only solution that spans from the data center for training deep narrow notch and running physically accurate simulations to a full stack in car solutions. Scaling from (inaudible) to level 5 fully autonomous functionality. Autonomous vehicle companies are harnessing this technology. Zeus recently-unveiled its level 5 bidirectional robotaxi powered by NVIDIA. Einride launched its next generation

cab-less autonomous truck using NVIDIA DRIVE Orin. And earlier this year, Mercedes announced a 56-inch wide MBUX hyper-screen powered by NVIDIA AI cockpit technology. This win builds on our momentum with Mercedes first generation MBUX system, which is now in 1.8 million cars.

We are in the early innings of a significant opportunity. We have built a multi-billion dollar design win pipeline for our self-driving AI cockpit solutions, which will drive a material inflection in revenue over the next few years. Our transformational partnership with Mercedes, announced last June, demonstrates the power of our evolving business model as we expand our addressable market and layer in software revenue. We are exceptionally well positioned to capitalize on the significant opportunity that lies ahead.

Moving to Data Center. Revenue was $1.9 billion, which exceeded our expectations, was comparable to last quarter and up 97% from the year-ago period, which did not include Mellanox. Data Center Compute revenue was up 45% year on year. Full-year Data Center Revenue rose 125% to a record $6.7 billion, including almost 70% growth from Data Center Compute. From a sequential perspective, the Data Center Compute's stronger than expected double-digit growth more than offset the anticipated decline in Mellanox revenue, which included a large non-reoccurring network sale to a single OEM in Q3.

Compute growth was led by Vertical Industries we're OEM partners continued ramping up their A100 based servers and our own DGX system sales were strong. Vertical Industries were well over 50% of strength in supercomputing, financial services, higher education and consumer internet verticals. Additionally, hyperscale customers continue to deploy the A100, driving both sequential growth and exceptionally strong year-on-year growth in Data Center Compute.

The A100 has been adopted by all major cloud customers globally and is being deployed by hyperscale customers for internal workloads. Still, we are in the early stages of adoption and expect continued growth this year. The ramp of the A100 has been smoother and accomplished by better visibility of the prior generation. Its universal AI training and inference capabilities as well as support for a wider set of applications and outstanding performance are driving [ph]five[ph] customer utilization, a clear sign of the A100's value.

Turning to Mellanox. We are seeing continued strong traction and robust momentum across our customer sets. Its revenue was up over 30% from Mellanox's Q4 revenue in calendar 2019 when it was still a stand-alone company. Year-on-year growth in the quarter was led by Hyperscale and large consumer internet customers which grew over 60% from last year, with several contributing record revenues. Consistent with our outlook, Mellanox had a sequential decline impacted by a non-recurring sale to a China OEM in Q3. We expect to return to sequential growth in Q1, driven by strong demand for our high-speed networking products, including the ramp of connect-X adapters with CSPs and all major server OEMs in their upcoming refresh.

We also see strong momentum in high-performance computing with HDR InfiniBand products. For example, we won six of the seven supercomputers awarded over the past

few months by Euro HPC. Starting next quarter, we will continue to provide color on networking as part of the Data Center market platform, but we will no longer break out Mellanox revenue separately. Looking forward, we are incredibly excited about the opportunities in Data Center. Accelerated computing is not only delivering super more laws gains in performance but is also an energy efficient and cost effective method of computing and virtually every industry is adopting technology with greater urgency as companies adapt to the new world of more distributed workers and customers.

As industries embark on this journey, they are also increasingly focused on combating climate change. To that end, the A100 performed AI computations with 120th the power consumption of CPUs. It powers Selene supercomputer which is number one on the Green500 list of the world's most efficient supercomputers. Indeed, NVIDIA's powered machines recently captured 25 of the top 30 spots on the Green500 list. Accelerated computing is not only serving the exponential growth in demand for compute, it can also help bend the power consumption curve. With accelerated computing, NVIDIA forward the computing industry.

Before I move to the P&L and outlook, let me give you an update on our proposed acquisition of Arm. In September, we announced plans to acquire Arm from SoftBank Group in a transaction that will create the premier computing company for the age of AI. At that time, we said it would take approximately 18 months to secure regulatory approvals in the US, the UK, the EU, China and other jurisdictions. Thorough reviews are typical with a deal of this size. This process is moving forward as expected. We are in constructive dialog with the relevant authorities and are confident that regulators will see the benefits to the entire tech ecosystem.

As we have said, this combination will spur competition. Together Arm and NVIDIA will provide greater choice to the data center ecosystem, a compelling alternative CPU architecture for the market and further enhance Arm's offering in mobile and embedded. Our intention is to increase investment in Arm's existing roadmap. Adding resources to stimulate growth in new markets, we love and intend to maintain Arm's open licensing model, a commitment guaranteed both by long-term legally binding contracts as well as our own interest in ensuring this investment is a profitable one for us. We are on the cusp of a new age in which AI fuels industries ranging from health care to scientific research to the environment. With this transaction, our vision is to boost Arm's potential so contrived in this new era and grow into promising new markets.

Moving to the rest of the P&L. Q4 GAAP gross margins were 63.1% and non-GAAP gross margins were 65.5%. GAAP gross margins declined year-on-year due to amortization of developed technology acquired from Mellanox, partially offset by product mix. The sequential increase was due to higher margins for gaming GPU and lower IP-related costs, partially offset by lower margin mix in our Data Center portfolio. Non-GAAP gross margins increased by 10 basis points year-on-year and was flat sequentially, in line with our expectations. Q4 GAAP EPS was $2.31, up 51% from a year earlier. Non-GAAP (technical difficulty) year ago. Q4 cash from operations was a record $2.07 billion.

With that, let me turn to the outlook for the first quarter of fiscal 2022. Revenue is expected to be $5.3 billion plus or minus 2%, with most of the sequential growth driven

by gaming. GAAP and non-GAAP gross margins are expected to be 63.8% and 66% respectively, plus or minus 50 basis points. GAAP and non-GAAP operating expenses are expected to be approximately $1.67 billion and $1.2 billion respectively.

For the full year, we expect to grow non-GAAP OpEx in the mid 20% range. GAAP and non-GAAP other income and expenses are both expected to be an expense of approximately $50 million. GAAP and non-GAAP tax rates are both expected to be 10%, plus or minus 1% excluding discrete items. Capital expenditures are expected to be approximately $300 million to $325 million. Further financial details are included in the CFO commentary and other information on our IR website.

In closing, let me highlight upcoming events for the financial community. We will be virtually attending the Raymond James Institutional Investors Conference on March 1st, the Morgan Stanley Technology Media and Telecom Conference on March 3rd and the Arete Virtual Semi conference on March 3rd. In addition, we will be hosting a virtual Investor Day on Monday April 12th, following the live stream of Janssen's opening keynote at our GPU Technology Conference. Our earnings call to discuss our first quarter is scheduled for Wednesday May 26th.

We will now open the call for questions. Operator, would you please poll for questions. Thank you.

## Questions And Answers

### Operator

(Operator Instructions) Your first question comes from the line of CJ Muse with Evercore ISI. Your line is open.

### Q - C.J. Muse

Yeah, good afternoon and thank you for taking the question. I guess, Jensen, a higher level question for you on the enterprise side. You're now a couple quarters into the ramp of A100 and curious if you could speak to whether you've seen any surprises here, any areas of specific strength worth calling out and any changes to how you're thinking about the size of this opportunity?

### A - Jensen Huang  {BIO 1782546 <GO>}

Yeah, thanks a lot, C.J. As you know, A100 is a very different type of GPU. This is our first universal computing GPU. It's great at high performance computing, it's great data analytics, it's great at training. And also for our highest end GPU is also the first time that it's incredible for infants. It's some 20 times faster than previous generation and it introduces some really, really exciting new computational formats like TF32, (inaudible) for training, and (inaudible) with a multi-instance GPU turning our GPU 1 GPU into a whole bunch of smaller GPUs, autonomous GPUs, to improve performance and reducing latency.

And so the capability is really quite exciting. We're seeing strength in hyperscalers as they continue to accelerate the adoption of AI. Some of the new applications we've spoken about a couple times before, the transition to deep learning for conversational AI speech recognition to natural language understanding, all the way to speech synthesis is now based on AI -- based on deep learning.

The other area that is growing incredibly fast is the deep learning recommender models. Just about everything that you do on the Internet is based on recommenders. There are hundreds of different recommenders out there, whether you're shopping or recommending music or recommending news or recommending search. And so all recommending ads. And so all of these different types of applications are driving growth.

For the first time, we installed our industrial application Industrial Data Center growing to be larger than hyperscale. And we're seeing industrial applications across scientific computing, where simulation-based approaches are now being infused with AI approaches for weather simulation, genomics, molecular dynamic simulation, quantum chemistry, even singularly quantum computing which is one of the really exciting areas.

We're seeing AI being deployed for big data analytic RAPIDS which is a NVIDIA has created open source platform for data analytics. Spark 3.0, which NVIDIA really led and as know GPU accelerated. So now, you could have big data in the cloud while doing big data analytics in the cloud on all the CSP (inaudible). And you could -- we're seeing a lot of excitement around financial services. Financial services and consumer Internet services are all really growing nicely. And so A100 is the production is just starting. We're going to see several -- a couple of years of continued growth yet ahead of us well as AI gets adopted in clouds and industries.

## Operator

Your next question comes from the line of Vivek Arya with BofA Securities. Your line is open.

## Q - Vivek Arya {BIO 6781604 <GO>}

Thanks for taking my question. Just a clarification and then a question for Jensen. On the clarification, Colette, I was hoping if you could give a little more color around Q1. Do you still expect Data Center to grow sequentially in Q1? I know you said that most of the growth was coming -- will come from gaming, but any color on the Data Center would be useful. And then Jensen, the question for you is in your press release, you use the phrase AI driving the smartphone moment for every industry. Could you help us quantify what that means? And where I'm going with that is, is there a number in terms of what percentage of servers are shipping today video accelerators and where can that ratio go over time? Is that a fair way of looking at the adoption of your technology and AI?

## A - Colette Kress {BIO 18297352 <GO>}

So, thank you, Vivek. Your question regarding the guidance as we lead into Q1. We had indicated that yes, a good percentage of our growth of between Q4 and Q1 both come from gaming. But we also do expect Data Center to grow. Most of our sequential growth

coming from gaming, but keep in mind, we also expect all of our market platforms would likely be able to grow quarter-over-quarter.

### A - Jensen Huang  {BIO 1782546 <GO>}

And I would say this of AI. The first phase of AI was new chips, the new systems, the use of the software, the new middleware, the new way of working, the new way of developing software, which the industry -- the world is now starting to call [ph]envel Ops. The way that software is developed and the way that it's deployed is completely different than the past. Something -- in fact, I heard really great term, software 2.0. And it makes a lot of sense. It's a computer that is writing software. The way that you develop software is completely different, the way you compute is different. And that was our first phase, and that started a journey that was some eight, nine years ago now.

The second phase was the adoption of using this in an industrial way for clouds. And and we saw it revolutionize new services, whether it's speech oriented services or search oriented services or recommender services, the way you shop, the way you use the Internet is completely different today. And so that's really the second phase. And those 2 phases are still continuing to grow, and you're still seeing the growth associated with that.

The third phase is the industrialization of AI. And some of the great examples, when I say kind of smartphone moment, I meant that it's a device with AI, it's autonomous and it's connected to a cloud service and it's continuously learning. So some of the exciting examples that I saw -- that I've seen and we're working with with companies all of the world, we have some 7,000 AI startups that we're working with and almost all of them are developing something like this. And large industrial companies, whether it's John Deere or Walmart, they're all developing applications kind of like this.

And basically it's an autonomous system, autonomous machine. In our case it's called Jetson. It's a robotics machine. That robotics machine was a card and it's called [ph]Dry and it's running an autonomous -- an AI application on top, and the AI is still on top. And this could be -- it could be moving device -- moving things around, it could be picking and placing, it could be just watching a warehouse and monitoring traffic and keeping traffic flow going. It could be connected to a car, and whenever the car -- whenever the fleet of cars needs to be retrained because of the new circumstance that was discovered, the cloud service would do the relearning and then would deploy them to -- into all of the autonomous devices.

And so in the future, we're seeing that these industries, whether you're in retail or logistics or transportation or you farming, [ph]bag[ph] tech, to lawnmowers, consumer lawnmowers. They're going to -- they're not going to just be products that you buy and use from that point forward, but it will likely be a connected device with an AI service that runs on top of it.

And so these industries are so excited about it because it gives them an opportunity to change the way that they interact with their customers. Rather than selling something once, they sell something and provide a service that's on top of it and they can stay

engaged with the customers. The customers could get a product that's improving all of the time.

It's just like your smartphone. And we saw what happened to the smartphone revolution and then we saw what happened to the smart microphone, the smart speaker evolution. And you're going to see the smart lawnmowers, smart tractors, smart air conditioners, smart elevators, smart buildings, smart warehouses. They will have autonomous capability, they'll all be driven by AI.

And so what's new for the industry therefore is that all of the enterprises in the world used to have computers for IT to facilitate to out [ph]hose their employees and in their supply chain. But in the future, only all of these industries, whether you're medical imaging or for -- or lawnmowers, you're going to have data centers that are hosting your products. Just like the CSPs.

And so that's a brand new industry and we have a platform that we call NGX which is the 5G Edge AI systems and we have the autonomous systems we call AGX which is what goes into (inaudible) and [ph]drawings. And between those two systems and the software stack that we have on top of it, we're in a great position to help these industries one at a time, transform their business model from the object-oriented business model, a thing based business model to a connected device discipline.

## Operator

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

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

Thanks for taking my question. First, I don't want to be pedantic I suppose, but I guess on the Q1 guide, you're say that gaming is the majority of the growth. Was that an absolute statement, or was that a percentage statement? Can you give us some idea of how you'd rank the sequential percentage growth of say Gaming versus Data Center versus other? Especially since it sounds like you've got $50 million in crypto specifics up that'll go into the other. And then I guess just briefly, could you give us some indication of where your supply situation and lead times are on your Ampere parts within Data Center? I think you'd said last quarter they were many months -- on six months plus. Are they still looking like that and is that the limiting factor at this point in terms of what you can actually ship on the compute side in Data Center?

## A - Jensen Huang  {BIO 1782546 <GO>}

Colette will take one and I'll take the other.

## A - Colette Kress  {BIO 18297352 <GO>}

Sure, let me start off of Stacy in terms of our guidance for Q1. As you know, we're still in the early innings of our Ampere architecture. Our ampere architecture as it relates to Gaming as well as what it relates to Data Center. As we articulated in our call, we have been really seeing continued uplift of folks adoption of A100 and it's going quite

smoothly than what we had seen in prior overall versions. So when we think about our guidance for Q1, there's many different types of conclusions that will happen at the end of the quarter in terms of what we ship. But all of our platforms can grow, but the majority of the growth from Q4 to Q1 will likely be Gaming.

## A - Jensen Huang  {BIO 1782546 <GO>}

Thank you. You asked a question about lead time. Our Company -- at the company level we're supply constrained. Our demand is greater than our supply. And -- however for Data Center, so long as the customers work closely with us and we do a good job planning between companies, there shouldn't be a supply issue for Data Centers. We just have to do a good job planning. And we have direct relationships with each one of the royalty fees and we have terrific relationships with all the (inaudible). And we could excellent planning between us. We shouldn't have a supply -- we shouldn't be supply constrained there. But at the company level, we're supply constrained, demand is greater than supply. And we have enough supply and we usually have enough supply to achieve better than the outlook, and we had that we had that situation in Q4. We expect that situation in Q1 and we have enough supply to grow through the year.

But supply is constrained and demand is really, really great. And so we'll just have to do a really good job planning. And meanwhile, one of the things that really came through for us, we have the world's best operations team. Our Company is really -- really has an amazing operations team. And we build the most complex products in the world, the most complex chips, the most complex packages, the most complex systems, and during Q4 they improved our cycle time.

And during Q1, I'm expecting them to improve our cycle time again. And we really are blessed to have such an amazing operations team. And so during these times, it really comes in handy. But overall at the company level, we're -- we expect demand to be greater than supply. We have enough supply to do those in the outlook and we have enough supply to grow each quarter throughout the year.

## Operator

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

## Q - Timothy Arcuri  {BIO 3824613 <GO>}

Hi, thanks. I had a question on crypto. I guess, Jensen, I know that the CMP stuff and the software drivers stuff that you're doing for the 36 and that's going to help a lot. But I think that there's like four or five of the big currency were going to move or at least are moving or they're on a path to move from the proof-of-work to a proof of stake, which is going to be a lot less compute-intensive. So I guess the question that I get a lot is, how do you assess the degree to which that drives GPUs back into the secondary market? Is there any way that you can get a handle on that? Thanks.

## A - Jensen Huang  {BIO 1782546 <GO>}

Yeah, if you look at the recent hash rates -- first of all the transition is going to take some time. It can't happen overnight, and so people have to build trust in the new versions and so will take a little bit of time. But I hope it does. I hope that people used our proof of stake over time and a little bit of these questions don't have to answered.

However, I don't have that much optimism either that it will be all proof of stake. I think that proof of work is a very general way of security -- securing the currency. And in the beginning, while any currency is buildings its reputation, it can take something like a proof of work to do so. And so I think proof of work is going to be around for a bit. And we developed CMP for this very reason, just so that there are different versions. We have different versions of our products for gaming, for professional visualization, (inaudible) computing, for deep learning. It stands to reason we have the ability to do a different version for CMP, and we can sell it directly the way that we go to market would be to go to directly to the industrial miners to -- and it's a great benefit to them so that they don't have to chase around spot markets. It's a great benefit to the gamers and because they want to game and the gaming demand is just incredible, it's off the chart. And so I think this is going to be really beneficial to everybody.

The recent hash rate growth was really a result of several dynamics. The first dynamic is the installed base. I mean most people thought that the -- once the mining the GPUs come back into the aftermarket, a small part does that. Some people do that. But the vast majority is they'll keep them, and the reason for that is because obviously they believe in the (inaudible). And so they're industrial miners, that's what they do. And so they keep it around for when the profitability returns and they could kick-start their mining gear.

We saw -- that's what we saw in the latter part of last year. We saw the hash rate starting to grow. Most of that was resulting of the installed miners. In the activating their -- it wasn't until earlier this year that we starting to see Dominion in our own GPUs. And when that starts to happen, there's several different dynamics. There's -- the primary source these days come from powerful ASICS and then there's some of that comes from our GPU and other GPUs in the marketplace.

And so I think that this is going to be a part of our business. It won't grow extremely large no matter what happens and the reason for that is because when it starts to grow large more ASICS comes in the market, which kind of mutes it. And when the market becomes smaller, it's harder for ASICS to sustain the R&D. And so -- and some of the spot miners, the industrial miners, come back and then we'll create the CMPs. And so we expect it to be kind of a -- to be a small part of our business as we go forward.

Now one of the important things is to realize that in the near term, because we're in the beginning parts of our Ampere ramp, we're only two quarters into it -- into a multi-year cycle. This is also the first time that we've completely changed computer graphics. RTX using ray tracing is completely different than (inaudible). And so this is a fundamental change in the way we do computer graphics and the results have been spectacular.

There's some 200 million installed base in desktops and 50 million in laptop. And the vast majority of them, we've only upgraded by approximately I think it's something like

50% at the installed base has been upgraded to RTX. And so there is a giant installed base and the installed base is growing that we need to upgrade to the next generation of computer graphics.

## Operator

Your next question comes from the line of John Pitzer with Credit Suisse. Your line is open.

### Q - John Pitzer  {BIO 1541792 <GO>}

Yeah, hi guys thanks for let me ask the question. I want to go back to Data Center. You've been very kind over the last couple quarters to call out Mellanox, both when it was a positive driver and when it was a headwind. I'm kind of curious as you -- do you -- when you're looking at the fiscal first quarter, is there anything of distinction to mention around Mellanox versus core Data Center?

And I guess as a follow on, the key metric that a lot of investors are looking at is when does the core Data Center business year-over-year growth start to reaccelerate, and some of that is just simple math. You're just comping very hard compares from last year. But, Jensen, how would you think about Data Center year-over-year growth in the context of a reopening trade or of any sort of new applications out there? What happened -- what helped the last time around was the move to natural language AI. Is there another big AI application we should be thinking about as we think about Data Center growth reaccelerating?

### A - Jensen Huang  {BIO 1782546 <GO>}

Yeah, we're expecting -- Mellanox was down this last quarter, and our Compute business grew double digit and offset -- more than offset the decline in Mellanox. We expect Q1 to be a growth quarter for Mellanox, and we expect to this coming year to be quite an exciting year for Mellanox. The business is growing and Internet, of course CSPs is growing, and InfiniBand for high performance computing on the switch -- the switches are growing. Switch business grew 50% year-over-year. And so we're seeing really terrific growth there.

One of the new initiatives, and we're going to see success towards the second half because the level of adoption, the number of engagements of them is on the BlueField DPUs. It's used for virtualization for hyperscalers. It's also used for security. As you know quite well, the future of computing in cloud and its multi-tenant cloud and there's no VPN front door to the cloud. You've got millions of people who are using every aspect of your computing so you get -- you need to have distributed firewalls and you can't have it just in one place. The intense focus of security across all of the data centers around the world is really creating a great condition for BlueField which is really perfect for them. And so I expect our Mellanox networking business to grow very nicely this year. And we expect Q1 to be a great growth quarter for Compute as well as Mellanox.

The tiller -- the great driving application for AI are several. Last year you're actually right that it was natural language understanding. And the transformer model and what is the core of BERT and other versions like that really, really made it possible for us to enable all

kinds of new applications. So you're going to see natural language understanding do text completion and it's going to be integrated. I think it was just announced today that it was going to be integrated into Microsoft Word. We've been working with them on that for some time.

And so there's some really exciting applications, but the new ones came -- that emerged recently are deep learning based conversational AI, Where the -- when the ASR, the speech recognition, as well as the synthesis are now based on deep learning. It wasn't before. And they were based on models that ran on CPUs, but now with these with these deep learning models, the accuracy is much, much higher, and it has the ability to also mimic your voice and be a lot more natural. And so the ability to be -- but these models are much more complex, they're much larger.

The other big huge driver is recommenders. This is something really worthwhile to take a look at. It's called deep learning recommender models. And recommenders have historically, whether it's through shopping in -- or personalizing websites or personalizing your store, recommending a basket, recommending your music, historically it's been in use that traditional machine learning algorithms. But because of the accuracy and the -- just the extraordinary economic impact that comes from an incremental 1% in accuracy for most of the world's large internet businesses, people are moving, very rapidly to deep learning based models.

And these models are gigantic, they're utterly gigantic. And this is an area that is really driving (inaudible). But we -- I expect us to see a lot of momentum. And the last one is the one that he was thinking of, which has to do with industrial 5G and edge IoT type of applications for all of the different industries, whether it's retail or whether it's just things for transportation, agriculture or warehouses to factories. And so we're going to see AI and robotics in a very large number of applications and industries and we're just seeing so much excitement there.

## Operator

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

## Q - Aaron Rakers {BIO 6649630 <GO>}

Yeah, thanks for taking the questions. I wanted to go back again on the Data Center business. You just mentioned, Jensen, that the BlueField 2 product poised to kind of ramp and materialize in the back half of the calendar year. How do you see that, is it an attach rate? I think there's some discussions in the past about all servers could potentially over time incorporate this layer of acceleration. How quickly should we think about that ramp? And then the second question is, can you just on a high level talk about how CPU -- how a CPU strategy you're thinking about that in the context of the broader Data Center market?

## A - Jensen Huang {BIO 1782546 <GO>}

Sure. If I just -- if I could just work backwards. And I believe that every single Data Center node will be outfitted with a DPU someday. And that someday is probably call it five years

from now. And we've been -- the fundamental driver of it it's going to be security. Every single application in the Data Center and every single node in the Data Center has to be individually secured. Zero trust computing, zero or confidential computing and zero trust computing, these initiatives are going to cause every Data Center to have every single application and every single node be secure.

Which means every one of those computers have to have a control plane that is isolated from the application plane. And all the applications cannot share the same resources because that application could be (inaudible), that application could be an intruder. No application could have access to the control room, and yet today, the software defined Data Centers, the software defined networking, software defined storage, all of the security engines are running in the same processors as the applications and that hasn't changed.

You see the cloud CSPs in the world are moving in this direction. Every single Data Center will have to move in that direction. So every node will be a DPU process for the software -- for the infrastructure. You're essentially going to (inaudible), the Data Center infrastructure be offloaded from the application plane. And it will be something like a BlueField.

So I think this is our next multi-billion dollar opportunity, the CPUs. We support every CPU in the world and we're the only accelerated computing platform that accelerates every CPU. Ironically, the only CPU we don't accelerate for AI is Arm. But we want to change that. Arm has such an exciting future because the nature of their business model and the nature of the architecture is perfect for the future of Hyperscalers and centers.

You want the most energy efficiency and in every single Data Center because every Data Center is power constrained. We are going to be powered constrained in every aspect of computing going forward. And so we would like -- love to build around the Arm processor and invest in building a great ecosystem around it. So that all the world's peripherals and all the world's applications can work and work on any one of the CPUs that we know today. And so we're going to start with high performance computing and start with AI and all of areas that we have a lot of expertise in, to build out that plant for them.

So you're starting to see one industry leader after another embrace Arm and I think that's terrific. But now we've got to energize it with all of the ecosystem support. It can't just be vertical applications, but we want to create a broad general Arm ecosystem.

## Operator

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

### Q - Mark Lipacis {BIO 2380059 <GO>}

Hi, thanks for taking my question. A question for Jensen I think. Jensen, if you look at the past computing eras, typically it's one ecosystem that captures 80% of the value of that computing era and mainframes is IBM and mini computers was dead, PCs Wintel,

cellphones Nokia and then Apple. So if you don't get the ecosystem right then you're splitting 20% of the market with a handful of players. So in this next era of computing, parallel processing or AI, I think you've articulated the most compelling architectural vision of the data center of the future with data center scale computing devices with CPUs, GPUs, DPUs integrated in the same box serving all workloads and imagine a virtualized environment. Can you help us understand where is the market in embracing that vision and where it's NVIDIA in building out that -- the ecosystem for that Data Center scale computing vision? And then maybe as part of that, to what extent is CUDAA the kernel for that ecosystem? Thank you.

## A - Jensen Huang  {BIO 1782546 <GO>}

Yeah, we are -- we -- I think we've done a great job building out the platforms for several ecosystems around the world. And the domains that we do incredibly well at are the domains that have to do with accelerated computing. We pioneered this approach and we brought it to high performance computing at first and we accelerated scientific computing and we democratized supercomputing for all researchers. Anybody who wants to have a supercomputer now can. And computing will simply not be the obstacle to somebody's discovery. We did the same for artificial intelligence, we did the same for visualization. We brought -- we expanded the reach of gaming tremendously. GeForce today is the largest gaming platform. It's the largest single largest body of computers that are used for gaming and in each case, we expanded the market tremendously.

We would like to do the same for Data Center scale computing as it applies to virtualizing these applications. These applications are also in the process. They've historically required dedicated systems, but they are moving into into a virtualized Data Center environment. And we are best at doing that. They run on a platform today, we have the ability to virtualize it and put it into the Data Center and make it remotely available. And so these applications -- these domains are some of the most important domains in the world. And so we're in the process of doing that.

By doing so and making our architecture available to CSPs and OEMs, we could create this accelerated computing platform available to everybody, and so that's -- you're seeing our journey doing them. First, creating and architecting this platform and then putting it literally into every single Data Center in the world. But we would also like to -- the next step of our journey is there's the phase 3 of AI, and it has to do about -- it has to do with turning every endpoint into a Data Center, whether it's a 5G tower, a warehouse, a retail store, a self-driving car, a self-driving truck, these are going to be -- they're all going to be essentially autonomous Data Centers. And they're going to run AI, but they're going to lot more and you're going to do -- they're going to do security and (inaudible).

It's networking is going to be incredible and it's going to run software 5G and GPU accelerated 5G we call Arial. And so these platforms are going to become Data Centers, they'll be secure the software is protected and you can't tamper with it. If you tamper with it the cores won't run. And so the capability of these clouds will move all the way out to the niche and we're in the best best position to be able to do that.

So I think in the -- in this new world of post (inaudible) law, post (inaudible) in this new world where AI and software, the right software, in this new world where we're -- Data Centers are on going to be literally everywhere and they're unprotected. There's no giant building with a whole bunch of people that secure it. And in this new world where software is going to enable the autonomous future, I think we are perfectly positioned for it.

## Operator

This is all the time we have for Q&A today. I will now turn the call back to CEO, Jensen Huang.

## A - Jensen Huang  {BIO 1782546 <GO>}

Thanks for joining us today. Q4 capped a truly breakout year for NVIDIA. The two biggest engines of our business, Gaming and Data Center, posted powerful growth. Gaming has become the world's largest media and entertainment industry and will grow to be much larger. And the game gamers will create and will play, they'll learn, they'll connect, the medium gaming can host any type of game and eventually evolve into countless metaverses, some for play, some for work.

Gaming is simultaneously a great technology and a great business driver for our company. This year, we also closed our Mellanox acquisition and successfully united amazing talent of our companies. Combined, we possess deep expertise in all aspects of computing and networking to drive the architecture of modern Data Centers. Cloud computing and hyperscalers have transformed the Data Center into the new unit of computing. Chips and servers are just elements of the Data Center scale computers now. With our expertise in AI computing, (inaudible) accelerated computing, our deep network to computing expertise and cloud to edge platforms NVIDIA is helping to drive a great computer industry transformation.

In our planned acquisition of Arm, the world's most popular and energy efficient CPU company, will help position NVIDIA to lead in the age of AI. This year was extraordinary. The pandemic will pass, but the world has been changed forever. Technology adoption is accelerating across every industry. Companies and products seem to be more remote and autonomous. This will drive Data Centers, AI and robotics. This underlines the accelerated adoption of NVIDIA's technology. The urgency to digitize, automate and accelerate innovation has never been higher. We are ready. We look forward to updating you on our progress next quarter. Thanks a lot.

## Operator

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

FINAL

Bloomberg Transcript