Morgan Stanley 2020 Cloud Secular Winners Virtual Conference

Company Participants

- Colette Kress, Executive Vice President and Chief Financial Officer

Other Participants

- Joseph Moore, Analyst, Morgan Stanley

Presentation

Joseph Moore {BIO 17644779 <GO>}

Okay, welcome back everyone. This is Joe Moore from Morgan Stanley, and I'm very happy to have Colette Kress, the Chief Financial Officer of NVIDIA. Before we start, I just quickly need to read a disclosure. For important disclosures, please see the Morgan Stanley research disclosure website at www.morganstanley.com / research disclosures. Please note that this call is from Morgan Stanley, institutional clients and financial advisors only. This call is not for members of the press. If you are a member of the press, please disconnect and reach out separately. Comments made on this call are not retribution by members of the press. Please note also this call is not for individual Wealth Management clients. If you have questions, please reach out to your Morgan Stanley financial advisor for more information.

So Colette, thank you very much for joining us today. We're really excited to have you. Maybe you could just start out, COVID has presented some unique challenges in the environment for all of your customers and suppliers, you guys have navigated it pretty well. Can you just start with maybe how you guys have navigated this so well, and how you guys are managing disruptions in supply chain and the health of your employees and things like that?

Colette Kress {BIO 18297352 <GO>}

Sure. Thanks so much for the question. We just released our results for our Q1 results last week, and we had provided overall Q2 guidance. Certainly COVID-19 was included in our Q1 results, when we announced our guidance for Q1 back in February. I think probably one of the things that in looking back was helpful for us, is the amount of our business, amount of our supply chain, which is in the Asia Pac area that allowed us to be front and center of the early stages of COVID-19 and what the early signalling was done. It's hard to know, with specifics how much was impacted with overall COVID-19, but some of the pieces that we saw right off the bat was a overall close of the retail channel, which impacted our gaming business at the start in terms of retail sales, as well as iCafe was now shut down in the overall Asia-pac

area. But what we quickly saw was a move to e-tail, People in the Asia Pac area had begun working from home, the schools were shut down, and it was a time for them to also learn from home and teach from home during that period. As you know, being in a shutdown worldwide, it led to all sorts of determining what the overall entertainment that would happen over that period of time and gaming really was a place that people turn toward, to find entertainment in their overall homes. So, we were able to quickly and agile company focus on moving towards e-tail focused on how to get the overall gaming to our end customers, and that worked pretty well. But then it moved to a pandemic, that pandemic within the quarter moved to worldwide and we were able to really leverage the learnings that we had in the Asia Pac and China area to think about that influence worldwide.

Keep in mind though, things in every single country, every culture was different. And it really focused us on understanding our suppliers, understanding our supply chain process and pretty much a connection on a weekly basis, daily basis in some cases to assure that we understood what they were facing. The overall Asia Pacific area restarted faster than some of the other parts of the world. But it was about getting overall people into the manufacturing and into actually building the overall products, there are still worse and challenges in terms of logistics. There were challenges in terms of just moving supply from place to place, and our connections with our top suppliers and connections with those manufacturing aided us during this process.

What was unique is our demand, particularly for our overall compute, and our data center stayed intact from what we thought at the very beginning of the quarter. And so really, we were just working through the overall supply and logistics issues through the quarter. We feel fortunate as a Company in terms of the businesses that we have chosen. And we know that there are many out there that are still challenged by COVID-19 with their overall businesses. We expect this to also continue through to Q2, we had highlighted that there are probably some continuations of COVID-19 into our Q2 overall guidance, particularly regarding the automotive business.

The automotive business and our legacy infotainment systems will be challenged and we will likely see automotive decline approximately 40% from where we had in Q1. We also expect our professional visualization project -- products to continue to seek where the overall business comes back from COVID-19 but will likely have a decline sequentially between Q1 and Q2 as well.

Overall gaming, gaming is still an area where I think demand is important, as well as our overall data center as we continue to ramp our new products such as A100 and the lineup for overall AI.

Joseph Moore {BIO 17644779 <GO>}

Great, that's a helpful overview and and it's an impressive job of maneuvering in what's obviously been a really tough environment for everyone. So I wanted to walk through the segments a little bit more in detail, starting with cloud, since that's sort of the theme of our conference, as well as the fact that it's now over half of the

Company are on a trajectory to be over half of the Company at least. So and maybe we could start just looking back a little bit and putting this strength that you've seen into context, you grew 80% year-on-year this quarter, which is not something I saw coming. But again last year, you saw this kind of digestion. So I want to talk about some of the newer drivers and the things that are driving the strength. But I think, you know, what's the post mortem on the digestion that you saw, a year and a half ago? And you know, the sort of prognosis for going forward for you know, is this going to continue to be a lumpy business? Or those kinds of things, you can just kind of maybe put that into perspective for us, that'd be helpful.

Colette Kress {BIO 18297352 <GO>}

Sure. So when we look back over, even the last three plus years with the introduction of the V100 and deploying that into so many overall data centers for internal use, as well as for the cloud, both a lot of lessons learned, but also a lot of overall expansion in terms of the customer starts and what put into place. When we look in terms of our initial rollout of V100, Sure, We started with the overall hyperscales. We started with the large customers as they found the importance of AI for their overall business models, for their overall monetization and really saw it as an important piece of the future of whether they think about their applications and even some of the basics of how they think about Search and switch on the Internet.

So, we began a significant amount of work working with them as they build out using overall accelerated computing. Moore's Law was quickly coming to an end. It was very well discussed and there was also a need to now allow AI to be used in the overall cloud environment. So, those three things around there really started an overall surge of interest, a surge of overall demand in terms of not only for internal engineering teams, but also surfacing that up to the cloud. Why they were important is that was the way that the overall enterprises, the higher education research could get their first hands on using overall accelerated computing with that.

Now, if you recall during that time, we also had some challenges with other providers having a shortness of overall supply. CPU shortage was being signaled in the overall market. So, when we look back -- overall digestion which can be calling when you're building out overall data center builds, the influence of also CPU shortages made it even more cloudy for people to understand where the overall demand would be. And through that time, purchasing ahead of what they need to do -- data centers, so they would not be short of any overall products is likely what happened over that period of time.

So, as we've turned now to a new architecture that we have announced for A100 and our overall work over the last three years, we have built out both a stronger list of customers, a (inaudible) of customers not only from a procurement and certain parts of overall engineering, but through and through understanding their workloads, their project's, so that we can actually help influence the importance of lead times and how long it will take in order for them to build out their overall data centers and or qualify and then what we can do to overall help that. That overall significant breadth and depth that we now have with our customers, we hope will help us as we go forward. But, there is still that likely, that from time to time a single customer may

be digesting, may be finishing out their build before they start a new one. Our hopes is that more overall large projects stuck up on top of each other that keeps the overall outlook smooth. And we believe that our process of working with them from a forecasting in that lead time will help improve, so that we have a better understanding of forecast going forward.

Joseph Moore {BIO 17644779 <GO>}

Okay, that's very helpful. Thank you. So maybe if we could talk about some of those new applications, first conversational AI. It's something that in the US are talking about and then interestingly, we started hearing it directly from the cloud customers sort of talking about the priority around these transformers for language translation and when we look into the models, the order -- the amount of complexity, the number of parameters is literally millions of times what the computer visions models looked like. So, clearly a big jump in complexity, pretty big compute workload, can you talk about where we are in that ramp. And is that going to be something that's narrow across a couple of cloud customers or how does that broaden out from an application set, so that, it's more interesting to a broader base of enterprises.

Colette Kress {BIO 18297352 <GO>}

Sure. So the concept of conversational AI and the very initial models that being built over last summer or late spring at about this time last year really started to surface. The BERT model surfaced, there has been many derivatives of the BERT, other ones that have been started by many of the overall hyperscales and with our direct working with them, we really understood their desire to build the overall complexity of models to support overall conversational AI. Conversational AI, the underpinnings of it, which is the natural language understanding which has many parts to it. It's not just a single overall function of work. There is a 3-step four-step process that is necessary for much of the natural language understanding that is happening in terms of understanding the overall speech, the tongue, the overall dialog, as well as understanding the words that we use in terms of, what is the question, what are you trying to solve for. And then lastly, being able to respond in a solid low amount of time in order for it to be within the reason of what you were expecting a response back.Now, what you're seeing is, we're still in the very early stages, but what they needed in order to complete that was, a size and ability of compute in order to help them through that process. So, our coming out with overall software, but more importantly, A100 allows them to expand their work to take on larger and larger models and more complex models to do their work. What we're seeing is the desire to not just bring in overall speech or to bring in just written overall data, but taking in any type of form factor and ability for the models to be at the larger sizes to address some of this very difficult work.

And so we're in the early stages, it's not necessarily something that's just for the overall hyperscales because you will see a lot of it also the part of overall cloud computing and that cloud computing is really about enterprises using the overall cloud and you can think of many other instances that they will need the overall conversational AI, whether that be overall call centers, whether that be ways that they converse as an overall company as a whole.

So, focusing first in terms of some of the work that the hyperscalers have done, that they really see this as a new way of collaboration that will likely happen in the overall enterprises as well.

Joseph Moore {BIO 17644779 <GO>}

Great, thank you. And the other area that you guys have talked about quite a bit recently, recommendator engines are various terminology to refer to those, the complexity of those models and sort of maybe less intuitive to me. But, obviously really important to the economics of your cloud customers. Can you talk a little bit about that work load?

Colette Kress {BIO 18297352 <GO>}

Sure. Direct mandate of engine, we see all across the overall Internet, not just in terms of the hyperscalers but also the Tier 2 or the consumer Internet companies that may be on top of overall cloud instances in terms of how they run their business, as you know just yet the basics that the news that comes to you every single day is generally a high and overall recommend data engine. What type of news interest you, what types of things have you read, how do we continue to provide things that meet your needs and overall interest, but this can go all the way in terms of overall procurement as well. When you think about what you procure as a business and-or personally, this is another area where recommendator engines are very important. You can think about how complex those models can become in terms of targeting marketing to you and also the speed back in terms of those recommendations, in terms of whether or not you are driving or whether or not you're just on your phone. So we are, with our overall processors, definitely able to process the significant amount of recommendator engines, but also remember, we have a set of application work that we are doing to better serve the overall recommendator engines in terms of software that supports those functions as well.

Joseph Moore {BIO 17644779 <GO>}

Okay, great. And then the other thing that, that I found pretty exciting about your business is AI inference, you know the NVIDIA chips have historically dominated the market for training, for building large databases but inference, the access of those databases has happened more on traditional CPUs, but that business has become pretty sizable for you guys. And, clearly you continue to make progress and ultimately as a market that that should offer similar size is training. So, can you talk about your progress there?

Colette Kress {BIO 18297352 <GO>}

Yeah so inference as we had outlined a couple of years ago, as we felt it was a very large market. A large market that we felt was appropriate for overall GPU computing. The overall inferencing we're talking about though is not necessarily the inferencing of the past. You're correct, it's a very CPU dominated overall piece. A lot of that has been single function or very binary type of overall inference, for the CPU again, may be perfectly fine to do. But, when you think about using GPUs for overall inferencing

going forward, the complexity of the inferencing and the models and the multiple stages puts the GPU in a great place both for the overall programmability, but also the overall speed that can be accomplished using overall GPOs. Using conversational AI past when you think of the training process for conversational AI and moving to the inference is a great overall use case of GPUs for overall inferencing.

We started this business with unique form factors for influencing that could be slotted directly into existing OEM servers in terms of a PCI-E slot. Our T4 for example is just well engineered in a candy bar size as well as very low wattage, in order for them to just add that overall infrencing to their overall compute and we have reached a point where inferencing is a sizable percentage of our overall data center business. It is solidly in the double digits percent of our overall data center. Now, what's interesting, it's in a doubling as year-over-year in this last quarter. But, also when you think about A100, A100 has been engineered and built so that it can accomplish both training and inferencing as we move forward. So, overall customer does not have to choose that says which workload am I trying going to, do you have the ability to choose A100 to do both. It has the ability for it to be partitioned into multiple virtual instances that each individual instance could be used for overall inferencing, and overall performance improvement in for versus using V100 but also just the flexibility of a single platform to accomplish both the training and the inferencing.

Joseph Moore {BIO 17644779 <GO>}

Great. And that's a good segway, I did want to ask about A100 in the Ampere family. These things don't come along that frequently for you guys. It's your first new architecture in 3 years. And it appears to be both from what you've said and what we've heard from your customers, significant advancement of state of the art. Can you talk a little bit about Ampere, the performance benefits that it brings? And, really what that will do for your business in terms of growing the number of workloads that you can address and driving your business going forward as well as dealing with competition.

Colette Kress {BIO 18297352 <GO>}

Great. So, yes, our A100 and our Ampere architecture, right before earnings, a week before was an opportunity for us to launch for the world to see what we've been working on for the last 3 years. Originally, probably targeted to be communicated our GTC San Jose back in March, but we just had to move that out a little bit and to try and find the right forum to communicate. So, we launched it just a couple of weeks ago, it is in full production and was a part of our overall Q1 results. The A100 architecture is our largest fleet in production -- performance versus any other architecture in the past. The A100 alone is 20X stronger in performance versus the overall V100. It is on 7 nanometers. It is the greatest performance on 7 nanometer. It is also the largest chip, on overall 7 nanometer. We're using this opportunity to launch A100 in a unique manner that we think will help both deployment and ability to bring it up in so many of our different customers. We are launching it with 8 GPUs together on an overall base board. That overall configuration, easily slipped in to

their existing overall infrastructure and will in the future also be with overall OEM service as well.

What is unique about it as we discussed, is it provides the ability to execute both training as well as inferencing. This is an important piece as much time is spent by customers, whether they be hyperscales, whether they be enterprises in terms of choosing their AI platform or the accelerated platform, and this is an important piece that allows the mobility to do both and create over 7 instances per GPU and being there is 8 GPUs, you have a multiplication factor of how many overall instances you can have. Now, only important in terms of the A100 architecture is the overall performance straight out of the overall hardware, but keep in mind, you could look at this as a very important software play that we do as well. We are on CUDA 11 now, that has come out. CUDA 11 is the underlying development language platform that we have on every one of our GPUs, as well as therefore our overall architecture with Ampere.

It is important piece as it is the underlying driver of many of the enterprise specific overall application that we will address. We also provide software that is specific to many of the key industries, vertical industries and key places that we believe accelerated will bridge to. Some of the things that we also launched with the Ampere architecture is new software enhancements that can help both for conversational AI, can help for overall recommendators, can help for overall anything in terms of the data analytics that is out there and the work with overall Spark. So our work, this is the 3rd generation Tensor quarters that we have with the overall Ampere architecture that is an important piece that allows both for high performance computing, but also the dialing up in terms of precision that you may use the overall GPUs for the positions of your different workloads.

So, we're excited about not only the overall Ampere architecture, but all of the software that we have also brought to market on top of overall CUDA and CUDA-X

Joseph Moore {BIO 17644779 <GO>}

Great. And then maybe that's a good segue to the competitive question. And in particular, there was a question on the webcast that I get a lot, which is how do you think about custom silicon from your hyperscale customers, obviously Google most prominently, but the others have talked about that as well. Can you just talk about the importance of the new product in that competition and do you expect that to be a negative for your growth at some point down the road, your customers are trying to develop customers (inaudible)?

Colette Kress {BIO 18297352 <GO>}

There are some cases and have been for a while. We are looking at custom ability and custom silicone largely due to large workloads that they may see in front of them that may have a overall cost in silicon to support that. What we found is, we are the platform that is probably the most agnostic to any platform that is out there that we need to work with, any CPU, any other types of components in the data center as well as any other type of software that is out there. So, our ability to work seamlessly

across different frameworks for AI as well as different types of cloud infrastructures is extremely important and why people think about using overall NVIDIA's GPUs to do that, you always have the ability with the frameworks to continue to write on top of those frameworks. However if the frameworks still need revision, still made more work to be done, you always have the ability to jump into the overall CUDA development platform to add and expand to that.

We believe our performance as well as meeting the needs of our customers is always been upheld quite strongly. If you think back to the overall fall timeframe when (inaudible) came out both for overall training as well as for inferencing, this is really a time for us to demonstrate from a benchmarking standpoint the hardware and software combination that we have, puts us in best of breed, and often when you see the custom overall ASICs that may be out there, it becomes very challenging for them to overall compute with the overall universal capabilities of the overall GPU and based on the speed of AI in AI development something may work today, but as you've seen the 3,000x improvement in terms of model size, as well as complexity of AI having that flexibility with our platform really enables us.

So, if our overall hyperscales or other partners look for a need that they see a workload that may need a specific ASIC, that may work for a time period, but we just don't think it really applies to the broad market that is out there, that really needs the flexibility as things are moving so fast.

Joseph Moore {BIO 17644779 <GO>}

Great. And for what it's worth, I mean, we've, we've heard that directly from some of the developers of custom silicon is that there are applications where it works and there is still very high demand for GPUs (inaudible), another question from the webcast is on the geography of the cloud business, can you talk a little bit about your positioning in China versus the US and is it possible that to the extent that the US goes through a digestion phase that China would continue to be strong for you.

Colette Kress {BIO 18297352 <GO>}

Yeah. So each of the regions does operate differently and we do serve all of the major overall cloud providers, around the world. The cloud providers that are here in the US are built with larger overall CapEx budget and are just bigger in nature than the ones that are in China. China has hyperscales but also has a second tier that is quite broad. And that's not just in China, but it is in the overall Asia Pac area and so, we continue to serve that, I think that helps us in terms of the differentiation of our overall customers in the overall cloud and we continue to see those projects overall stack up. All go through some form of digestion. But, the more ability to expand our overall customer such projects at each one of them can follow if any single one customer goes through digestion.

Joseph Moore {BIO 17644779 <GO>}

Okay. So maybe just sum all this up on the, your organic cloud business, you've sort of said, it seems like a pretty healthy growth in July. And you said that you probably

need more time to assess what happens beyond July, which seems totally reasonable. But, it seems like a business that offers a lot more breadth of application and breadth of customer than what you saw a couple of years ago, is that all fair?

Colette Kress {BIO 18297352 <GO>}

It is fair. When you think about A100 we're in the early stages of a multi-year ramp. The demand is strong for A100. We do have that visibility into Q2 which what we used for providing our overall guidance for Q2. As we look at Q3 and Q4, just given the overall environment out there, we're just need a little bit more time to really size, in terms of what we see for Q3 and Q4 and we'll provide that when we get to that time.

Joseph Moore {BIO 17644779 <GO>}

Great. Makes sense. Then maybe talk a little bit about Mellanox, you've talked about the rationale for doing the deal, their business has obviously been really strong. So, now that the deal is closed, can you give us an update on how this fits into the NVIDIA portfolio and maybe, I know you acquired Cumulus right after, it seems to be that you're committed to networking and maybe a bigger way than I realized before that second transaction, it seems like this is an area where NVIDIA plans to continue to grow.

Colette Kress {BIO 18297352 <GO>}

Yeah, so our Mellanox acquisition, we're really pleased that the acquisition did close about a month ago and our Q2 results will have a full quarter of undrawn Mellanox business. Mellanox was doing quite well in this last year. I think you saw, even after the overall announcement of the acquisition and our continued partnership with them, they still stand-alone did quite quite well over this last year. So, we're really pleased to bring them on part of our overall business and will be incorporated into our data center business.

Now the question is why, why is that a great addition to our family of products that we have in our overall data center business. As you've seen us focus in terms of accelerated computing, accelerated computing, one of the key areas is overall AI, but it's also our base of overall high performance computing and overall supercomputing. Mellanox brings their expertise absolutely into those same areas of customers that we do. What we've learned is not only compute and the time that is spent overall computing data information is essential in terms of that acceleration, but also the other components are very, very key to influence overall acceleration of the overall data center as a whole.

When you think about the modernization of data centers, as we go forward, really thinking about the other layers that software defined and acceleration can influence acceleration as a whole is where we're overall focused on. Their leadership in terms of in the interconnects their leadership in high-performance computing and overall supercomputing puts them as probably one of the best partnerships that we could imagine, bringing on board at this time for data center. You'll see us both working

together with the same sets of customers, but also thinking about products as we know, understand the full footprint of the data centers with these customers on how we can build better things to influence acceleration going forward.

So right now, just up merge of who we sell to, getting that right product to market is where we're focused, but you'll see more and more of a partnership, both at the software level as well as definitely the hardware level of new products that we can bring into market. Now Cumulus. Cumulus we announced that we are in process of overall purchasing, it's not closed yet, but this is another example of software defined, software defined in terms of the overall networking that can be a key component in terms of us spreading our wings of all of the different processes within the data center.

Joseph Moore {BIO 17644779 <GO>}

Great, thank you. So, we could spend this whole time on Just cloud and I didn't get to ask you about TJX or edge based computing, which I know, are important initiatives. But I do want to get some gaming questions in, you're gaming numbers have been good, better than I would have expected, because if you told me the challenges that you've alluded to with traditional retail, with Internet cafes, the supply chain there, is there a case to be made that there is pent-up demand from those things and the Internet cafes open back up, that there's more incremental demand or has there been a shift where in the regions where those cafes are prominent that people have shifted more to gaming from home?

Colette Kress {BIO 18297352 <GO>}

Yeah, thanks for the question regarding the success that we've seen in overall gaming through this period. Remaining agile during this period has really benefited us, as we thought about all of the different ways that we sell gaming, to make sure we could address the different regions around the world, the different countries in the world, as they worked through their overall COVID-19 issues. The iCafe's right now have probably returned to business, but not necessarily at 100%, as you could imagine, social distancing and overall density in the overall iCafe's as they function is so important. So, they are all carefully bringing them back up on board and we are watching to see how we can help them in terms of get back into business because it is such an important part of their overall economy as well when they think about that.

The retail is the same thing, how quickly can the overall retail come up to speed. It's still unknown and it's not clearly up to anywhere near 100% at this time, there is a long way to go. But, reaching the overall gamers whether it be high end overall laptop, whether it be a overall desktop card that they can buy in e-tail and, and self-build their overall desktops at well at home and or ability for them to share doing a high-end overall compute with what they need to both work and or learn in the overall home has really benefited us. We also saw demand surge as it related to overall steam, the amount of hours in time that they were playing online. This took the opportunity for folks to jump on GFN which you know is our cloud service that we recently just rolled out in February as an overall subscription offering as well. We've seen a surge not in each region, being very differently as they use a streaming

service to get there for sometimes their first taste at overall gaming as they have an owned device to do that. So, is there pent-up demand, is there more just too early to overall say that, the demand is solid, based on what we've seen, which doesn't surprise us. We've always found that during some of the hardest economic times, overall gaming and gaming on a PC platform is actually a great way to entertain and actually can be quite economical because there's so many different offerings at different price points to support that. We didn't talk about it, but also our console business or console business with Nintendo is doing quite well and we're here to support Nintendo in the demand that they are seeing for the for their great console switch that they have right now.

Joseph Moore {BIO 17644779 <GO>}

Great. Talking about ray tracing. I mean it's hard to pioneer new stuff. There is always sort of the chicken and egg issues of the complexity of writing software for something new and it's kind of groundbreaking is real time ray tracing. Can you talk about where we are in that, it seems like you've seen an inflection where the demand is, has really picked up there and you -- I know you can't talk about next generation silicon products. But, I know historically when you've had these kinds of breakthroughs like shader algorithms, the next-generation that implemented, that much better ones, it's already become something that's in the ecosystem has been pretty good. So, can you just put some perspective around the importance of ray tracing in your line up. That'd be great.

Colette Kress {BIO 18297352 <GO>}

Yeah, ray tracing is an important piece that will be the next generation of graphics as we go forward, not only for overall consumer graphics and therefore gaming graphics, but also for the enterprise. So, it's making a material impact on both of those major markets. What we're seeing is all of the great publishers around the world have jumped on overall ray tracing, it was a little bit of a chicken and the egg that shows what comes first. The overall software or the overall hardware. We introduced it both starting back with Microsoft, in the overall development on top of the overall APIs as they became available. So, for us to come into market as the leader and first-to-market has really, really breached a tremendous following of all of the different publishers, but also those that are building overall game into the future.

We have right now quite a star lineup of terms of games that are on and using ray tracing, we have 33 RTx games that have been announced that are actually shipping, probably most importantly is the best selling game out there, Minecraft is really a transformational shift for them. If you're familiar with overall Minecraft and it's overall architecture and the use of overall ray tracing to bring realistic type of pictures. It has been a really transformation for that overall game. But, keep in mind, there is a tremendous long list of other superior cache of games out there. Warrior 5, Macquarie 5 is out there, Call of Duty, Control, you going to see Wolfenstein, battlefield fire, we can go on and on, cyber punk will probably be one that comes out in the future as well. So these, as you can see are important games and I'm sure every publisher as it thinks about games for the overall holiday season or beginning to write games knows that they have to make that decision on how to incorporate

overall ray tracing, because in the future, it will probably be a very important for high-end types of graphics games.

Now, being the leader into market and having the overall hardware and the performance that supports it, is only one piece of it. Keep in mind, we have added significant amount of software to enhance the overall ability for ray tracing such as deal assess. Deal assess is an important piece of using AI in overall gaming, what that does is it takes the features of high end ray tracing and breaks that down to where you can infer in between 2 pixels what that overall completion of the overall frame should be. It improves the overall performance on ray tracing that gives you that realistic and great high end overall graphics as well.

We'll continue to take this not only to gaming, but to the enterprise, so many applications in film industry or things that are in terms of CAD drawings or otherwise, the use of realistic materials, we are going to get close in a closure that we will look at a catalog, we will look at a set of pictures to try and determine, is that real or is it rendered. But, there is a great wave right now, we're excited in just such a short amount of period of time, the uptake to support high end ray tracing. So, we look -- shouldn't look at this is ended, we have more to come in terms of the software enhancements that we have as well as just the overall architecture that supports ray tracing.

Joseph Moore {BIO 17644779 <GO>}

Great. Sounds good. I can't wait to see what what developers do going forward. So, maybe you could talk about the financials a little bit. Starting with the buyback, you suspended the buyback while you were working on the regulatory approvals for Mellanox, you still have quite a bit of authorization, when might we see you renew that and what's the criteria that you're thinking about there?

Colette Kress {BIO 18297352 <GO>}

Sure. So we used the time in terms of exposure to Mellanox to try and begin the first half of rebuilding our overall cash, we had enough cash on hand as you know to complete the overall Mellanox acquisition with cash, but we wanted to make sure we had a solid amount of reserves on hand to support the investments into our business or in business in terms of a capital perspective, investments capacity and or any future overall M&A. So, we are building back that overall cash portfolio, we stay overall nimble and agile as we think about overall buyback, you are correct that we do have a authorization with our Board to repurchase up to $7.4 billion of stock repurchases before the end of 2022.

So, we will stay nimble to find the opportunity, once we continue to build back our net cash position and then get started, in terms of stock repurchases.

Joseph Moore {BIO 17644779 <GO>}

Great, thank you. And then gross margins continue to be moving upward, 66% is a level that I didn't you get through this quickly, is that largely mix driven? And are you seeing change and kind of like-for-like margins as you move forward? Will, the main driver be the move to HPC cloud?

Colette Kress {BIO 18297352 <GO>}

Yes. So on our gross margin, our Q1 results were also a record level in terms of gross margin. And as we move to Q2 and our guidance for 66, it would also be also a record quarter. Mix definitely is the largest driver of our overall gross margins. When we refer to that as mix, it's really talking about platforms that are encompassing a significant amount of software. The software, as you know is being built with our overall R&D engineering teams and therefore the overall cost of writing software is in our overall OpEx. It's therefore influenced our ability to maintain great gross margins for most of our overall businesses. We'll continue to find opportunities to enhance our overall gross margins going forward. I do believe that our past growth in gross margins has been quite, quite strong over the last 5 years or 10 years. And as we go forward, we'll make some movements as well, but maybe not as the speed of what we've seen over the last 10 years.

But, what you should see in terms of something like overall data center, we're in the early round of A100, when you're in the early ramp and you're also in the early days, in terms of the manufacturing process of that, we'll probably see yields improve a bit. We'll look in terms of value engineering types of projects across our platforms, including the overall A100 architecture. From season to season or quarter to quarter, we will have different mixes across our products, overall, but also a different mix within some of our overall market platform such as gaming.

When we move to higher end platforms in such as gaming, we can achieve overall higher gross margins as well. We haven't mastered perfectly yet, but we have all of our gaming platforms at the high end. We do see people coming into market for the first time, by our higher end platforms, but we still have volume SKUs that bring first time gamers or even those that are coming back to get the overall new architecture. Ray tracing has allowed us to provide higher end platforms and therefore achieve higher gross margins as well. So, a lot of different mix factors, but yes, our focus of ours to continue going forward to improve our gross margins.

Joseph Moore {BIO 17644779 <GO>}

Great. And maybe we could finish with a question from the webcast, which kind of ties everything together is your automotive business. First, can you touch on the declines this quarter, is that a macro issue, or is there any share issues in the infotainment space and then probably more importantly, when do you start to see the autonomous programs and the big investments that you've made there start to come to fruition from the revenue side?

Colette Kress {BIO 18297352 <GO>}

Yes. This last quarter and the quarters we go into Q2, have been some of the most challenging quarters for the automotive industry. The overall crisis around COVID-19 again was an industry that was quite impacted, quite difficult for them during this time, we are here to support them as they bring back up their manufacturing lines and as they work through the overall sales of their cars. But, before we even entered COVID-19, there was a lot of discussion with auto manufacturers as they knew that the future related to AV and EV were very important part of their strategy as auto manufacturers going forward.

Looking for a solid way to both develop on these platforms, but synergize across possibly a single platform that could take them from ADAS all the way up to robotaxis, became an important topic for them, supporting multiple platforms in-market or during development is tricky for them and very hard to find the resources to support multiple platforms. We've been working with them on our end-to-end platform that can support ADAS all the way up to robotaxis to help influence a more strategic and refined investment that they need to make to develop these overall platforms. So, as we work with them, we're excited to have an end-to-end platform today that can already go into production, but also working with them for the long term as they think about their -- as their fleet entering us. So, stay tuned as they work through our COVID-19 bringing things back. But again, we probably are seeing something more in the 2 years after term, before we get to solid amount of production of autonomous vehicles in market.

Joseph Moore {BIO 17644779 <GO>}

Great, well Colette, thank you very much for spending time with us today and sharing all this information, this has been great.

Colette Kress {BIO 18297352 <GO>}

Okay, thanks a lot

Joseph Moore {BIO 17644779 <GO>}

Thanks. Bye.