Daniel Newman: Sarah Peach, Samsung Semiconductor, welcome to the 2022 Six Five Summit. Great to have you here.

Sarah Peach: Thank you, Daniel. Pleasure to be here.

Daniel Newman: Yeah, it's great. I love talking about data. I think this is one of the topics that just about every enterprise, organization, institution, has front and center. It's basically going to drive the future of our businesses, of our world. And so talking to people like yourself that are leading product and innovation at companies, it's fascinating to me. So first and foremost, the processing, the speed of data, it's happening so fast. The way we handle data is going to be the difference between the haves and the have-nots. Give me some trends that you're seeing right now in the storage and in the data space.

Sarah Peach: Yeah, absolutely. Yeah, I think the trends are basically more, more and more. We're seeing more data volume. We keep generating, here as humans, we keep generating more and more data. By 2025, we're expecting 400 exabytes of data each day, which is a phenomenal volume. And that's both an opportunity and a challenge for our customers, figuring out how to deal with those massive quantities of data and extracting value from it. We're also seeing data created in new locations, from traditionally enterprise, and more recently cloud, and now IoT edge data generation. It's moved from industrial and consumer devices into, for example, cars, autonomous driving. That's terabytes per vehicle, per day, that you're seeing there. So new locations for data generation. And we're also seeing more data intensive applications like AI and ML that require huge amounts of storage and memory. So it's really more, more and more.

Daniel Newman: Yeah, I was just thinking about that. So if I'm hearing you, and I think this definitely matches up with my research, is pretty much everything that we're doing is generating volumes of data that are exponential compared to what we were dealing with in the past. And of course, in order to get things like AI and ML right, to get autonomous driving, to get maximum productivity from applications, it means we need all that data to be accessible for training, for learning, for intelligence and for insights. What are the challenges that you see being created through all this volume of data?

Sarah Peach: Well, you are absolutely right. The challenge is using that data. As it grows, the larger the data gets, the harder it is to move it around. There's this concept of data gravity, that data weighs something, is heavy and difficult to move around. And that's absolutely true. As your data grows, it becomes more costly to move it from one location to another, or from storage into compute. It takes longer to move. That slows everything down. You need more bandwidth. And it also takes more power, literally more energy to move bits from one location to another. So these are really challenges that you deal with when you start to get these very large data volumes.

Daniel Newman: So when I speak to people in roles like yours, one of the things that I always look for is, you're talking to these customers, I've been to the Samsung Semiconductor Tech Days, I've seen the ecosystem, the partners, customers, users of your technology. How are you in terms of, in the innovation, in the new products that you're building and developing, how are you solving these problems? I think that has to be top of mind in every conversation that you're having.

Sarah Peach: Yeah, absolutely right. I think our customers are seeing these limitations of existing compute architectures and looking for solutions. How are we going to handle these large data volumes more efficiently? And we are taking, I'm from the Samsung Semiconductor business, so that's memory and storage. And so we are taking a memory and storage focused look at how we can help customers. And one way in which we're approaching this is figuring out how we can do more within memory or within storage. So how can we add functionality to our memory or storage devices that will help our customers deal with large volumes of data? And one way in which we're approaching that is adding processing power to our storage and our memory devices. So we've been doing that both in memory with acceleration dim, so we are adding logic to DRM modules in memory, also high bandwidth memory. We have processing in memory that provides AI acceleration, so that's on the memory side.

And then on the storage side, we're adding compute into solid state drives. That's our SmartSSD computational storage drive. And both of these approaches are really giving our customers more options for working with their data, more options for accelerating results from that data, more options for optimizing their workflows to get more out of their data.

Daniel Newman: We're definitely seeing a continued pressure on memory technologies to keep up with the number of cores, the amount of compute, and everybody knows that all the cores in the world without memory are pretty much useless. So, there's this tug of war between compute and memory, you're seeing different fabrics and technologies being introduced. How do you plug in more memory? Well, most parts are of the stack are composable. Memory presents a number of different challenges. I think what you mentioned with computational storage is interesting. Talk a little bit about how that's changing the compute paradigm.

Sarah Peach: Right. Yeah. I think computational storage is bringing high performance compute to traditional storage devices. So what that means is that you can process data and analyze it directly on the drive where it's stored. You don't have to pull that data from storage into memory, into the CPU of your system, in order to get some insight from it. This is not actually a new idea. It's been around for a long time, but all of this growth of data is really making it more valuable. Now people are realizing, "Gosh, we have too much data to be able to shove it all into the CPU complex." And so what's new is that there is availability of devices, so some practical hardware that now supports computational storage. We launched ours in January of last year, 2021. And we've been learning since then about, what are the best applications? What are the true benefits of allowing processing in storage?

Daniel Newman: What are you finding? What are those benefits?

Sarah Peach: Mostly the benefits are what I would say TCO driven, what our customers are looking for is lower total cost of ownership and computational storage helps them in a couple of different ways. It means faster time to insight. If you are processing data where it's stored, you can get results faster. So if you're a data analyst, you are running more queries, you're getting those query results faster. It's also more efficient and overall reduces the load that you have on your other compute, your main compute. So you can either save on main compute, not buy such an

expensive server, or you can use that freed up main compute for other applications. So it's really an overall improved application performance and better use of your infrastructure.

Daniel Newman: Yeah. That makes sense. One of the ways that I always look at a technology, whether it's brand new or it's been in the market a while, is the ecosystem. How well adopted is it? As I mentioned, you're very embedded in an ecosystem because memory technology, storage technology, compute technology, networking, all the fabrics that enable us to do anything from high performance computing all the way down to just every day things on our mobile devices all requires these things to interact efficiently and effectively. So as it relates to computational storage, is this being adopted? Who's using this? How fast is it being rolled out? Where is it being picked up in the markets now?

Sarah Peach: Okay. Well, that's a great question. Let me just comment on your ecosystem point because that's a very, very key point. As I mentioned at the beginning, we at Samsung, we're adding functionality to memory and storage, and that means that it's no longer such a simple, straightforward product that we're offering. So we are actually working much more closely with ecosystem partners to bring these types of products to market, software partners, hardware partners, channel partners, even standards bodies are all in the mix to get products like this out to the market and make sure that they're useful and usable by customers. So, you're absolutely right, the ecosystem is key for these types of products.

Regarding computational storage and where it is in the market at the moment, we are still definitely early stages in the market for computational storage. Customers are becoming more aware both across the OEMs, across hyperscalers, more aware of what's possible with computational storage. There's some venture funding, startups active in the market, new applications emerging all the time, but it's a very dynamic situation. So I think we're in the thin end of the wedge here with computational storage and plenty of room for growth.

Daniel Newman: Yeah. I think that seeing that adoption and seeing more and more of the market starting to embed this technology, to include this technology, is going to be the validation. And so you're already starting to see it. It sounds like there's a lot of interest right now, as we're trying to solve problems everywhere from supply chain to technology refreshes, to digital transformation, to hybrid work, all these things that are going on concurrently, basically compute is an underpinning of all of it. And so I look at, when I say compute, it's everything in that entire ecosystem of compute. So it'll be good to watch how this computational storage picks up, gains momentum, gets more market coverage, because it's, like I said, it doesn't sound like it's new. It sounds like it's one of these things that's going to dramatically improve things that are already, it's both innovative and iterative in terms of helping move things forward.

So with that in mind, what about specific applications? Where would you tell users that are looking at this and saying, "Okay, where does this fit into our strategic utilization of resources?" Where does computational storage fit in? What are some of those best killer applications for the technology?

Sarah Peach: Right. So at a very high level, I would say computational storage is great if you are looking for things in very large data sets. So, at the highest level, that's basically what computational storage can do really, really well. If you've got a huge data set, you know there's some gold in there and computational storage is going to allow you to find that gold very quickly and then do analytics later in more detail, just on the good stuff. So dramatically faster search through large volumes of data. What does that mean in terms of actual applications? Database acceleration is a great example. Scanning and filtering through databases just to focus on the subset of data that's particularly relevant for your analysis is a great use case for computational storage. And a second use case that I particularly love is cybersecurity. One key aspect of cybersecurity is searching through log and event data.

That's a critical part of protecting your infrastructure. But as data grows, the tools for ingesting this log data, this event data and scanning through the data, don't really scale very well. Tools like Splunk, for example. And using computational storage means that you can continue to use these tools, these software tools with much, much faster results, even as your data scales up. So cyber is definitely a good example of a situation where you are looking for tiny things in large data sets and where time matters. If you can search, for example, a search that would take half a day or a day is now taking 25 minutes with computational storage. That means you are finding these breaches faster, which means you are resolving, remediating those breaches faster, which means the overall cost of that breach is going to be substantially lower. So I think that's a great application of computational storage.

Daniel Newman: Yeah. That makes a lot of sense. I love anything that moves cybersecurity forward. When you just see the volumes of attempted breaches and the impact that it's having right now on businesses, companies, it's like you're always playing defense. And so any technology that speeds that up, and you mentioned Splunk, who also spoke here at this event, it's definitely something that they're focused on. You guys also have other partners that you are working closely with, AMD joined us here and Xilinx. FPGAs are red hot right now as well. I understand that there is a bit of a partnership brewing there with Samsung.

Sarah Peach: Absolutely, more than brewing. It's well established. Yes.

Daniel Newman: Talk about that. Tell me about how the two organizations are working together.

Sarah Peach: Yeah. So we've been working very, very closely with Xilinx, now AMD, right from the beginning and through the launch of this product, the compute component in our SmartSSD is a Xilinx FPGA. And so this has been a joint go-to-market. I mentioned ecosystem earlier. You did too, right? They are an absolute critical partner for us in getting the product out. If you want to buy a SmartSSD, please go to Xilinx for that. And what's been great is a combination of Samsung and AMD's experience. AMD, obviously, well previously Xilinx, really understanding the acceleration market and Samsung really understanding, of course, the storage market. And so a combination of those two views is really valuable.

Daniel Newman: Yeah, that makes a lot of sense. So as we come to time here on this conversation, I always like to get a little bit of a look into the future through the lens of those developing, building,

accelerating and innovating. So, what do you see as the future? How does computational storage go forward? Are there other applications that you see driving here? What's next for this technology?

Sarah Peach: Well, I'd love to see more applications in AI and ML. That's obviously a massive market and I think where computational storage can really help customers. Also, edge applications, computational storage is effectively a whole system in a box, storage, memory, and compute in one small form factor. So I think there's plenty of room for growth there. And I'd also love to see, since we're early in the market, I'd love to see computational storage instances available in the public cloud because I think we're at a stage where we need to see how people are going to use this. We've identified a couple of great use cases, but there are many, many more that we are not aware of. And so I think just general availability of computational storage is really going to help to get the word out and folks to really understand how best they can use that to help them in their business.

Daniel Newman: Yeah. Sarah, I really appreciate you joining here. If I could summate a lot of this, it's bringing compute and storage closer together is a huge opportunity. And with the exponential data scale that we were all seeing happen, the only way that organizations are going to be able to utilize all that data is by enabling it to be accessed more seamlessly, more rapidly. So it's getting rid of that latency as much as possible. It's getting the two working harmoniously and it's people like you that are involved in developing, building and rolling out those solutions. And I really enjoyed having you here today to talk about this.

Sarah Peach: Thanks, Daniel, for the opportunity.

Daniel Newman: Yeah. Let's talk more about this soon and thank you so much, Sarah, for joining me here at this year's Six Five Summit.

Sarah Peach: Thanks.