Patrick Moorhead: Rob, it's great to see you, and thank you so much for kicking off the data track, data and AI track for The Six Five Summit 2022.

Rob Bearden: Well, thanks, Pat. It's great to be here.

Patrick Moorhead: I have followed Cloudera for years now, and it's really great to see the industry impact you've had with CDP, otherwise known as Cloud Data Platform. Rob, for those who aren't familiar with your company like I am, how would you summarize or describe what you do?

Rob Bearden: Sure. Well, today, Cloudera is focused on optimizing a hybrid data strategy that will enable businesses to take full advantage of all their data everywhere. Last year, we became a private company again because it enabled us to accelerate innovation to meet the needs of our customers more effectively and, actually, grow the business faster. We actually crossed an important milestone last year by achieving more than a billion dollars in revenue. And given the success of our CDP platform and the demand we're seeing around the world, we expect that we're going to double that billion-dollar number in a much shorter period of time. We're very, very excited about where we stand and where we are in the industry today.

Patrick Moorhead: Yeah, Rob, I've been super impressed with the amount of data that customers are using on your platforms, and I think you're squarely in the big trends of hybrid cloud, multi-cloud, and in the future of SaaS. You've been doing a lot of things at the same time. I know some of the largest enterprises are your customers. Can you talk about how they're working with data and the challenges that they face because you're right at the epicenter of all of this.

Rob Bearden: Yeah, absolutely. I speak to customers almost every day, and one common theme that they're still struggling with is how to manage and analyze the huge volumes of data that they continue to collect and have to manage. And so they look to us as they embrace hybrid infrastructures and new architectures like data lakehouses, data fabric, and data mesh because this is what they need in order to solve their data challenges. And this is especially true when it comes to leveraging AI and machine learning. And our core focus is hybrid data, and hybrid data is everywhere. It's not just in a single cloud or just in the data center. It starts in legacy on-premises. It then moves to private cloud. Of course, it's in multiple public clouds, and it's at the edge streaming in from third-party services and partner applications. And it's actually in our customers' products and in the customer interactions that they have and all the employee activities that they monitor, and it just keeps exploding. And they look to us to help them solve those problems and take advantage of those data opportunities.

Patrick Moorhead: Rob, I am super happy where we've come as an industry. It's funny, if you had a conversation with somebody maybe 10 years ago, there may have been a board mandate that said, "Get everything to the public cloud. We're not too sure why we want to do that, but let's get it all there." And I just knew, from being around IT for over 30 years, that just didn't make sense. I appreciated the vigor at which it was discussed, but we're finally here. And I would say, four

years ago, the industry finally all agrees now that it's a hybrid world out there, so I am certainly happy that you headed that direction a few years ago. What I'd like to talk about, though, is how does Cloudera help customers make sense of all that data? We see these slides of petabytes of data used. We're doubling the amount of data every week or every month. We've all sat in front of this, but you're actually working with customers to help make sense of all that data. How do you help them manage that?

Rob Bearden: Yeah, well, the question we work to answer with all our customers is: How do you take advantage of all your data across your entire enterprise and get value from it in the fastest way possible? And from our perspective, to accomplish that, you need to be able to manage the full life cycle of data analytics and machine learning across one common platform. And with hybrid data, you have to be able to apply data analytics across workloads and across all form factors, public, private, and the edge.

Patrick Moorhead: Yeah. I forgot that little edge part, that little edge part that's going to grow faster than anything else. Sometimes, as technologists, we are really focused on explaining all of this stuff to IT or the CIO, but, in the end, hybrid data and being able to wrangle that and have a strategy around that is all about delivering business value. How are you seeing your customers get business value out of the hybrid data cloud?

Rob Bearden: Well, you're spot on. Businesses have to get the most value from their data, and they need to get that value today. Not some distance in time and space, and that may or may not ever happen. And we focus on making it easier for the individuals, the data teams, and the entire enterprise to manage access and analyze all their data everywhere, irrespective of where it sits, public cloud, private cloud, or in the edge. And the fact that we do this for some of the largest enterprises in the world gives us a very unique perspective. It gives us insight into what are the best practices of developing a very data-driven culture and providing the employee base and the data scientists easy access to all the data that they need to drive the value back out of it.

And this is really important because where data flows, good ideas and better decisions always follow. But companies have to get there without adding layers and layers of complexity. And our customers continue to tell us that they don't want to have to choose between competing priorities and, in the end, end up with still a lot of silo data sets that they still can't access and bring the value back through.

Patrick Moorhead: Yeah, I totally agree. Whether it's increasing your revenue, trying to find a different business model, maybe as a service business model, trying to lower costs, increasing, improving customer NPS. It's funny how all roads lead back to being able to manage big amounts of data out there. So for some people watching, they may be like, "Hey, this is great. I love the theory, Rob and Pat, but how does this work in a real life scenario?"

Rob Bearden: Yeah. Well, let's talk about one of our customers who is actually one of the world's largest and premier luxury car manufacturer. And they just cannot compromise when it comes to the management and performance of their data analytics. The company has massive volumes of data streaming in from thousands of fleets of cars all day, every day. And the data needs to be

analyzed in real-time for a variety of different use cases, everything from predictive maintenance to warranty analytics, to customer 360 views. And just to click on the predictive maintenance use case, the company is able to actually gather insights into what's going on across all the different components in the car in real-time. And they can really detect and predict: Is there going to be a component that's going to break or a part that's going to break?

And what that lets them do is reach out proactively to a customer, arrange to either go pick the car up or have them bring the car in. They make sure that they have the part and devices. They're ready when the car shows up. They repair it quickly, get the car back to the customer. And what that's done is, it's just driven customer satisfaction and the net promoter scores across every one of their customer sectors through the roof. And obviously, it translates into accelerated growth, great brand net promoter scores, and just great repeat loyal buyers. And that's just one of the many use cases that they've applied leveraging data to drive customer success.

Patrick Moorhead: Yeah. I love those types of proactive experiences. And I will admit, and maybe I overanalyze this just because I've been in product management and marketing, but I want to go back to those companies because I feel like they know me better. And then, when you look at the new generation, it's so funny, a new generation Millennials, and then Gen Z, that's their expectation, is that they're not looked at as a mass of a customer, but they're looked at as a unique person with unique needs, and very true for this car example that you gave.

Rob Bearden: Yeah, exactly. And it is a expectation that every enterprise has to deliver on or be dismissed. And this particular car manufacturer, they absolutely know, that if they don't get their data strategy right, everything they're trying to do to drive customer satisfaction and better margins is going to fail. There's just hundreds and hundreds of examples of companies that try to drive business value with data, and they're not successful because they don't get the data strategy right. Or they don't get the architecture right. And in today's world of data strategy, that will ultimately drive the cloud strategy, which, in turn, is going to drive the business strategy and new differentiated business matters. And this customer had multiple other top... Excuse me. Okay. And this customer, along with multiple other top car manufacturers, have chosen us to realize their business strategy through this modern data architecture approach around streaming. And they've been very, very successful driving customer satisfactions and better margins by leveraging it.

Patrick Moorhead: Yeah. So, Rob, driving business value from data sounds great, makes the top line, maybe the bottom line or the stickiness better, but we have data everywhere. It's literally everywhere. And I think you map that out. Whether it's the edge, whether it's the end point, whether it's at a colo, whether it's in the public cloud, whether it's in the private cloud, whether it's in legacy IT, it's everywhere. So what about security and compliance? Is this something that keeps your customers up at night?

Rob Bearden: Well, it's a table stake issue. And we work with companies in some of the most highly regulated industries in the world. And it's absolutely imperative that we provide the highest levels of security governance and metadata management. And we're central to not only their data

security strategy, but also, in how they're enabling the actual data movement in a secure way across multiple data centers throughout many countries and all the compliance requirements across those different companies. And then, we are obviously moving that data across public clouds and private clouds. One great example is one of the largest stock exchanges in the world. They rely on us to help them with real near-time decisions by analyzing a tremendous volume of data across multiple vectors and thousands of different data points. And they want that speed in decision making in real-time, but it also has to be coupled with the strictest of compliance and have a secure data exchange across each one of the form factors that they leverage those data sets from, and that's what they leverage CDP to do.

Patrick Moorhead: Yeah. When I first started researching Cloudera, one of the things I was most impressed about were the logos that are your customers. And first of all, they were processing a gigantic amount of data, but also, the risk of a breach, either on the security side or now on the privacy side, was immense and, at least from my point of view, the biggest part of your strategy was having a consistent security governance and metadata management across the entire data life cycle, all the way from data ingest, all the way to the end when you're doing machine learning inference. But, hey, on this hybrid data architecture thing, talk more about that. It seems like everybody is saying, "Hey, I'm hybrid. We're hybrid. You're hybrid. I've always been hybrid." How are you different? How is this different from what your competitors deliver?

Rob Bearden: Yeah, well, let's start with our view of hybrid. And our view of hybrid is the customers have to have the freedom to move existing and future data applications, the data itself, and the users bidirectionally between the data center and across multiple public clouds. And so, CDP and our data services are delivered through four primary components. The first is our data platform. And this is where we storm process all the data across private and public clouds through an identical experience. The next area is our data fabric. And this is our unified data governance layer that dynamically helps orchestrate disparate data sources intelligently and securely in a self-service manner. And then our data lakehouse enables the world's only open lakehouse, supported by Iceberg, that integrates and unifies the capabilities of data warehouses and data lakes and helps deliver multifunction analytics on the same data set at scale.

And with our data mesh, it helps the data product owners in the company and the data teams deliver data applications to the line of businesses with a low-code/no-code approach. And, Pat, from the differentiation point of view, Cloud Air is the only company that makes the hybrid data strategy a reality today. We can manage and store data across any form factors, public, private, or at the edge. And we provide a unified security and governance mesh that will manage the entire pipeline of data analytics in a consistent and unified manner. And we can do that at scale that no one else in the industry can today. And we have customers that are doing it really successfully.

Patrick Moorhead: Yeah, Rob, and as an industry analyst, I don't say this lightly, but I vouch for what you just said. I've looked at a lot of companies. A lot of companies do data in the public cloud really well. Some companies do it really well in the private cloud. Some have stove pipes of excellence that are hybrid, but only hit maybe one out of six elements of the hybrid strategy, so you have to find yourself cobbling together multiple vendors to do that. So, bravo on pulling that off.

So, hey, let me move to the next topic, and this is about partnerships. You have a very large partner ecosystem on the hardware side with the chip providers like Nvidia, Intel, and even the large hyperscalers like AWS and Microsoft Azure. And I'm curious, how does this help you or help you help your customers as it relates to, let's say, artificial intelligence and machine learning?

Rob Bearden: Yeah, sure. Well, Pat, those are, obviously, clearly some of our most strategic set of partners. But we work with an ecosystem of more than 2000 partners, and they're a big part of our success as they help us jointly deliver innovative solutions that our customers drive value with. And as the amount of data and demand for analytics and AI just continues to grow and explode, we now need GPU-enabled hardware and software that supports accelerated... Let me go back.

Patrick Moorhead: Yeah.

Rob Bearden: As the amount of data and demand for analytics and AI continue to grow and explode, we need GPU-enabled hardware and software that supports accelerated data workflows and real-time data streams. And the work we're doing with some of the chip vendors around advanced data engineering workflows and automating data pipelines is really game-changing. We partnered with Nvidia to provide native integration of GPU hardware for Spark 3.0 workloads on our private cloud deployment of CDP. There's also a bunch of other initiatives that we're focused on. An example of that is we're working with partners like LinkedIn, AWS, Netflix, and Tencent on building Apache Iceberg, which will enable our customers to bring the analytics tools of choice that they want on the platform rather than just being relegated to one single tool.

Patrick Moorhead: Yeah. So, Rob, I love the technology. I've been in and around accelerated computing for longer than I would admit. But the biggest question that I get is, what results do companies get out of these big, superpower, accelerated computing chips?

Rob Bearden: Yeah, well, using our solutions and our partner solutions, we commonly see customer realize a 10x performance improvement and a 3x cost savings right out of the box without additional resources or code changes. A great example that I would outline is IQVIA, who's one of our largest multinational companies. They serve the healthcare IT in clinical research space, and they were able to increase deployment speed for their data science and machine learning workflows across the business by 90%. And this has empowered multiple teams at IQVIA across their business to deliver enterprise-ready data clouds in their line of businesses to get value from in a matter of days and, instead of the old world, it took multiple months.

Patrick Moorhead: That's a really impressive results for IQVIA. I've also heard that you're working with Intel.

Rob Bearden: Yeah, absolutely. Intel is a longstanding partner, and we're working with them on enabling Cloudera on the Intel architectures like Xeon processors and isolate to dramatically optimize data processing workloads and enable faster insights and better security with a lot less complexity. We're also partnering with the hardware platforms like Dell, HP, and several others, to help us deliver a secure, interoperable and portable hybrid data platform with accelerated computing across any form factor. And the reason this is important is because it helps us and our partners

build solutions to deploy data in any form factor across any environment to drive data and streaming analytics solutions that are optimized for these chipsets and hardware platforms.

Patrick Moorhead: Yeah. I'm a big fan of accelerated computing, and I'm glad to see you partnering with these folks directly as opposed to maybe tangentially. So, Rob, what about the public cloud and services and solution companies? Who are you working with in this space?

Rob Bearden: Yeah, that's an important part of our strategy. And we're partnering, actually, with several of the hyperscalers and cloud providers, including, of course, IBM, Red Hat, AWS, Microsoft, and Google, to make it faster and easier for our customers to move their workloads and applications between the public and the private clouds, as well as be able to leverage the edge as well. And we partner with companies like Accenture and Kyndryl because they're actually the ones taking the ecosystem technologies and enabling the use cases and creating the business value for our combined customers. But the reality is, there's billions of dollars of prepaid cloud credits that many large enterprise customers are just sitting on. And Cloudera is the best company in the world to help them maximize value across any hyperscaler on any form factor, whether via the edge, private cloud, or public cloud.

Patrick Moorhead: Yeah. So, Rob, based on what you're saying, it's really clear to me that the underlying architecture is really a key for enterprises to get a handle on their data. Can you tell us more about when and how to use lakehouses, mesh, and fabric?

Rob Bearden: Yeah, let's unpack that a little bit. As big data evolved to cloud data, it's now transitioned over to hybrid data. And it's gotten more complex for companies to access and use and create value as it's distributed to so many different form factors. And that's where modern data architectures like data lakehouses and data fabric and data mesh come in. And these new architectures are engineered to handle the complexity that comes with that automatically so that the IT teams don't have to deal with it and to create that integration tax that comes with it.

And in the Cloudera world, we start with the data fabric, which is powered by Cloudera SDX, which orchestrates all the disparate data sources intelligently and securely in a self-service manner. And this gives our customers a unified, trusted, and very comprehensive view of all their data. And what's key here is that SDX allows you to do this across multiple public clouds and on the on-premise private cloud platforms and all the way out to the edge. We then, with our open data lakehouse, we enable the multifunction analytics on both streaming and store data in the cloud-native object stores across hybrid multi-clouds.

And as I mentioned earlier, we did this keeping open and interoperability as a design principle within our product set, which is why we're working in the community, using Apache Iceberg as our next generation table format, in order to enable our customers to use any analytic tool of their choice with the lakehouse. Then finally, our customers want to enable a data mesh, which lets them treat data as a product. And now, we give them the ability to own and serve up their data products, but do it in a consistently secured, governed, and orchestrated manner across multiple clouds and on-premise private clouds all the way out to the edge. So this hybrid deployment capability with centralized governance is absolutely fundamental to enabling the

data mesh architecture our customers want. And this is really going to be empowering the next generation of data applications our customers are looking to deliver back to their business.

Patrick Moorhead: Yeah, Rob, that was a really good, succinct description of how your technology works to set up customers for success. But to some, this could be daunting. It could be very complex. And I recognize, just by looking at the logos, you have some very smart customers with a lot of resources, but even for them, and even maybe for smaller companies, that this could seem daunting to them. So how do you remove the complexity for your customers in this process?

Rob Bearden: Sure. Well, of course, our goal is to get as many of our customers on a self-service journey as possible, to get value from the data as quickly as possible. But we also help enterprise determine: What are the deployment readiness? What is the deployment readiness for your different data sets and workloads across different form factors? We help them think through both from an architectural and a financial standpoint. Is it really pragmatic to put it in the cloud? Is it better served in a private cloud, or is it something that we can do as a real-time analytic out on the edge? And we help them build the business and technical implementation plans, taking into account the multiple vectors, including financial risk, the business value they're going to capture and the ability to implement within their objectives and timeframes.

A good example of that is what we did in our partnership with HelloFresh. For those of you who may or may not be familiar with HelloFresh, they're the largest meal-kit company in the world. And they had a very classic centralized data management strategy with a small number of internal and external data sources. And it was really the typical siloed approach to managing data. But they experienced a massive and extraordinary growth across the company, and they realized, as part of that, that data was going to be and had to be a key strategic asset that could provide a competitive advantage. And they partnered with us to build a data mesh that enabled them to really deliver a series of trusted data products back into the rest of the company.

And the day that data mesh has helped them eliminate all their data silos and provide multiple data products that have delivered significant value and differentiation back to their customers providing things like dashboards that monitor error rates, making sure that they have a 360-degree view of their customer's interaction with them, their customer success rate and providing dynamic recipe recommendations to make sure that they're managing and expanding the share wallet that they're getting from their customer. And our goal is to give every company that we work with the ability to achieve these same types of advantage to move faster in a much and easier way.

Patrick Moorhead: Yeah. So, Rob, we talked a little bit where Cloudera came from; The capabilities that you're providing customers today. And this year, what an exciting year. You're going from public to private, but let's talk about the future now. What does the future of data look like?

Rob Bearden: Yeah. Well, Pat, at the macro, very macro level, going from a public company to being a private company has allowed us to make longer-term decisions that accelerate innovation and growth for our customers and our partners and the community in a much faster pace. And this was quite different, obviously, than the quarterly cadence and constraints that we had to live in, in the

public world, but what we've also been able to really accelerate is our R&D investments. And this has allowed us to deliver on our cloud services and the tooling required to help legacy customers migrate from the old Hadoop platform to CDP public and private services and, actually, more than 50% of our customers have already migrated to CDP, and the remainder are on track to get there in the next 12 to 18 months.

And so, as a result, we've been able to really accelerate our growth and actually significantly improve our profitability. But because some of these investments that we've been able to make, one of our customers, who's actually one of the largest financial institutions in the world, has been able to accelerate and migrate their workloads over to CDP. And this particular bank has eight lines of business, and the bank's data science and analytics team was tasked with leveraging the bank's data assets and wanted to leverage the power of AI to deliver multiple use case across fraud management, compliance, credit scoring, and customer personalization. And they've initially deployed CDP private cloud to provide real-time data ingestion and leverage advanced GPUs and spark grids to support fast computations on massive amounts of data.

And this has enabled a very cost-effective data usage at scale set of use cases. And they've been able to reduce the time that had typically been required to implement new projects from six months to just a few weeks. And it's also decreased the total cost of their infrastructure by half. And keep in mind, this financial institution has 98,000 people in IT trying to deliver data solutions. And they turned to us to simplify the process of delivering data to their teams and businesses and how they quickly capture value back from that data.

Patrick Moorhead: That is impressive. Six months to a few weeks, cut infrastructure costs by 50%. That is amazing. They must be heroes over there. So, Rob, been a great conversation, and I'm going to ask you what I ask most of my speakers. What's the one thing that you would leave us with, some words of wisdom about Cloudera and data?

Rob Bearden: Well, great. Well, Pat, what I want to leave you with is that Cloudera is helping the world's largest enterprise customers solve their business problems by helping them with their data management, data analytics at scale. And the reason why we're so successful in doing that is because we provide customers with a robust data fabric, an open data lakehouse, and a modern data mesh that allows them to manage data anywhere in their enterprise at scale, and do that on any form factor, private cloud, public cloud, or at the edge. And most importantly, we do this through a unified security and governance plan with open cloud storage formats. And, Pat, I believe this is really just the beginning, and I think the next five years will be as exciting and transformative for the data industry as any five-year period we've seen over the last 20 or 30 years. And we're confident that Cloudera is going to play a significant and dominant role in shaping the future of how the enterprises capture value for managing their data in the hybrid world.

Patrick Moorhead: Yeah, Rob, I 100% think that we're headed toward, I mean, we're in a hybrid cloud world right now, and I'm super excited what you're bringing to the table with data lakehouse, fabric, and data mesh. And I'm looking forward to seeing what's next to Cloudera, but I appreciate you kicking off the data track for The Six Five Summit. Thank you so much.

Rob Bearden: Well, Pat, thank you for having us. We've enjoyed being here and look forward to the next tracks coming up. Have a great day.

Patrick Moorhead: Thanks, Rob.