

REPORT REPRINT

Cloudera adds user-specific toolsets with CDP Data Engineering

OCTOBER 30 2020

By Agatha Poon

The company has refreshed its Cloudera Data Platform portfolio with the launch of CDP Data Engineering on AWS. As the name suggests, the new offering aims to provide data engineers with a curated experience while managing resources and jobs across teams. Cloudera also highlighted demand trends in Asia-Pacific, and discussed growth potential in the post-COVID-19 era.

THIS REPORT, LICENSED TO CLOUDERA, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, WAS PUBLISHED AS PART OF OUR SYNDICATED MARKET INSIGHT SUBSCRIPTION SERVICE. IT SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR RE-POSTED, IN WHOLE OR IN PART, BY THE RECIPIENT WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



S&P Global Market Intelligence

Introduction

Following the GA of Cloudera Data Platform (CDP) Private Cloud, Cloudera has refreshed its portfolio of CDP with the launch of CDP Data Engineering. As the name suggests, the new offering aims to provide data engineers with a containerized, curated experience for scheduling, orchestration and managing data pipelines across teams at scale. The company also discussed its growth opportunities in Asia-Pacific as local companies seek new ways to adapt to the new normal.

451 TAKE

Cloudera has made good efforts in aligning its product capabilities with user-specific requirements. Being able to maintain a steady demand amid global lockdown is a good indication of the market's readiness. With the growing complexity of data management, some business organizations had already offloaded their responsibility to data specialists before the COVID-19 pandemic. This trend is likely to prevail as companies begin to reset their priorities for post-COVID-19 recovery. The company is aware of the geopolitical risk, and is in talks with potential partners for alternatives. Overall, Cloudera seems to be heading in the right direction, with CDP to accelerate sales in both developed and developing Asia.

Context

In Asia-Pacific, Cloudera already has strong partnerships with the likes of Dell, Accenture, IBM and Red Hat, and continues to deepen its product integration with cloud hyperscalers AWS and Microsoft Azure for seamless service experience. With AWS, Cloudera has been actively involved in a number of joint sales opportunities and co-marketing programs. The company also started working with the Google cloud team, although there are no joint offerings yet.

Given that CDP Private Cloud is running exclusively on Red Hat OpenShift, both parties are working collaboratively to drive thought leadership with executive roundtables and workshops. With growing tension between China and the US, selling into government agencies and state-owned enterprises could be a daunting task for any foreign vendor. Cloudera is no exception, and is devising its game plan. Thus far, the company is in talks with potential partners for strategic partnerships, yet no further details were made available.

While subscription services remain a primary revenue generator, accounting for appropriately 75% of the company's business, CDP Public Cloud is gathering momentum as Asia-Pacific businesses transition to the cloud. In terms of directional trends, customer requirements are evolving from managing data platforms to building and tuning application strategies for competitive differentiation, the company notes. Increasingly, businesses are more willing to offload data management and support to third-party data specialists.

An internet service provider in South Korea is one such example. Instead of running and maintaining its data platform, the company has prioritized its internal resources for building AI models and applications while offloading the responsibility of data platform management to Cloudera. As data management continues to grow in complexity, Cloudera is making a fresh push across the board to grow its share of what it thinks will be the next boom amid digital transformation.

REPORT REPRINT

The company reported \$214.3m in revenue for Q2 FY2021 ended July 31, 2020, up by 9% YoY. Its subscription revenue stood at \$191.5m for FY2021, up by 17% YoY. While it doesn't provide revenue breakdown by region, it claims to have experienced much stronger growth in Asia-Pacific, and expects to see a positive outlook for 2H FY2021. As discussed in our previous report, CDP Public Cloud has already caught the attention of business organizations in Asia-Pacific, and the COVID-19 pandemic has brought a new sense of urgency for companies to migrate workloads to the cloud if they haven't already done so.

Product

At the core of the company's data strategy is the Cloudera Data Platform, which is now running across major public clouds (AWS and Microsoft Azure) and the private cloud environment. Since the official debut in September 2019, the company has brought forth the concept of a 'curated experience' while providing toolsets that are specific to various stakeholders from within the data lifecycle.

Following the introduction of data warehouse and machine-learning services as PaaS for business analysts and data scientists, respectively, the company has added Data Engineering to CDP. Touted to be the only cloud-native service designed for data engineers, particularly those who are using Apache Spark for data processing, CDP Data Engineering is now available for customers running CDP on AWS, and will soon extend support for Microsoft Azure.

In a nutshell, Data Engineering serves as the 'backbone' to operationalize and debug data pipelines, through processing, curating and qualifying them, and then putting them into downstream systems. Prior to the Data Engineering service, Cloudera says, customers deployed Hadoop using Data Hub clusters running on VMs – which is a form factor, not a cloud-native deployment.

Now, with Data Engineering – which is delivered as a managed Apache Spark service on Kubernetes – customers can have the agility to scale resources up and down, and portability to move workloads between and across multiple clouds. Since Data Engineering is fully integrated with the Cloudera Shared Data Experience (SDX), all users have full visibility into the operational workflows, with unified data and metadata security and governance policies across hybrid and multi-cloud environments.

Apache airflow scheduling, which is a preferred tooling based on user feedback, has been integrated into CDP for creating pipelines using Python. There is a new tuning and visual troubleshooting interface to enable self-service support. Accordingly, the interface has been codified with the best practices created by the Cloudera team, based on years of experience in debugging and optimizing Spark jobs. With visual presentation, data engineers can easily monitor the status of resource consumption and utilization, identify bottlenecks, and perform troubleshooting in a timely fashion.

To simplify job management, customers can define the configurations of the job using a pre-loaded wizard with typical configurations to choose from, and then create jobs with a few clicks. In addition to Python, Data Engineering supports Scala and Java jobs. As far as data integration is concerned, data engineers are focused on automation capability. With Data Engineering, Cloudera enables API-driven pipeline automation. By using a consistent set of APIs that are accessible through a CLI and REST, data engineers can automatically set up and manage Spark jobs, and streamline integration with existing CI/CD workflows.

Demand trend

Although the COVID-19 pandemic has taken a toll on some industry verticals, such as manufacturing with the supply chain being disrupted and business slowing down, the company continues to see demand in the region. The public sector in Singapore is moving ahead with data/AI projects, as are the banking and government sectors in Australia. Local governments in China are in recovery mode, investing in IT infrastructure projects including AI, big data and cloud computing as a way to invigorate the economy, which has seemed to fare better than other nations in Q3.

Businesses in the ASEAN region are motivated by the need to have contingency planning as they come to terms with the new normal. For developing countries such as Indonesia that are undergoing their own digital transformation, Cloudera expects to see a positive 2H performance. India is another growth market for Cloudera, nearly doubling its sales YoY. Japan has always been synonymous with technological advancements in multiple business disciplines; ironically then, many local businesses have struggled with the WFH policy. While the corporate culture is partly to blame, local businesses are playing catchup when it comes to digitizing their business processes and investing in automation.

Competition

Given that Cloudera has extended its platform capability and is focused on delivering higher-value use cases such as machine learning, AI and advanced analytics, Cloudera believes that its new competition comes from the likes of IBM Watson and other data science platform vendors.

SAS, for example, is looking to exploit its expertise in machine learning and other data science realms with a cloud platform dubbed Viya. TIBCO is seeking to serve enterprise data science and analysis needs on-premises and in the cloud. When it comes to delivering data science and machine learning as a service, Google TensorFlow is gaining attention. H2O.ai is looking to make data science enterprise-ready using open source technology.

SWOT Analysis

STRENGTHS

With years of operational experience, Cloudera clearly has the expertise when it comes to providing user-specific toolsets.

WEAKNESSES

Adding new capabilities makes sense; however, it does require customers to buy into CDP to get the full benefits.

OPPORTUNITIES

The market for managed data platform as a service is likely to grow steadily as companies continue to reset their business priorities for post-COVID-19 recovery.

THREATS

The ongoing China-US trade war could evolve in ways that hinder Cloudera's ambition in the region, but the company is not alone when it comes to navigating through times of uncertainty.