

REPORT REPRINT

IBM combines databases and AI with updated and unified Db2 Hybrid Data Management Platform

JULY 8 2019

By Matt Aslett, James Curtis

The company has unified the codebase for the various products based on its Db2 database with version 11.5, which also includes improved support for artificial intelligence to enhance data management and make it easier to develop AI-based applications.

THIS REPORT, LICENSED TO IBM, 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.



Introduction

IBM has announced the launch of version 11.5 of its Db2 database, which includes built-in support for data science tools and frameworks for the development of artificial intelligence applications, as well as the infusion of AI functionality to improve the efficiency of data management. The launch of version 11.5 is also significant because it sees IBM move to a single SQL engine code base for the multiple products and services that share the Db2 brand, as well as others that are part of the wider Hybrid Data Management portfolio that also incorporate database functionality.

451 TAKE

IBM has a solid track record as a provider of database and data management functionality and has expanded its portfolio over the years through internal development and acquisitions. While the functionality has expanded, the portfolio had also fragmented, both in terms of branding and the underlying code base. IBM has taken steps in recent years to address this. Today, the combined Hybrid Data Management Platform is a differentiated offering that enables users to adopt multiple products as their needs change rather than having to license separate products or services. The infusion of AI functionality is in its early stages, but IBM has a strong roadmap to improve both Db2's automation capabilities driven by AI, as well as its support for data science and the development of AI-based applications.

Context

Having expanded its database portfolio through internal development as well as acquisitions over the course of several decades, IBM has been more focused in recent years on consolidation. First came the adoption of Db2 as a brand name in 2017 for the various on-premises, cloud and hosted database products and services. Then came the 2018 launch of the Hybrid Data Management Platform family, which includes the Db2, Db2 Warehouse, Db2 Hosted, Db2 on Cloud, and Db2 Warehouse on Cloud products and services, as well as the Db2 Big SQL SQL-on-Hadoop engine, the Db2 Event Store event processing and analytics environment, and also the Integrated Analytics System analytics appliance product.

With the launch of Db2 version 11.5, all of the various products and services in the Hybrid Data Management Platform family now include the Db2 Common SQL Engine, which provides a common code base for the underlying database engine that supports the various workloads, including transaction processing, analytic processing, big-data analytics and event processing.

In doing so, IBM has gathered together a number of capabilities that were previously spread across individual products and services in the Hybrid Data Management Platform family. In particular, this includes functionality such as queryable external tables, Netezza compatibility and columnar grooming that was previously only available with Db2 Warehouse/ Db2 Warehouse on Cloud.

AI database

In addition to introducing the Db2 Common SQL Engine, Db2 11.5 is being pitched as an 'AI database.' This means two things: that Db2 now features built-in support for data science and the development of AI-based applications, and that Db2 now includes automation capabilities driven by AI. In terms of built-in support for data science and the development of AI-based applications, Db2 11.5 includes drivers and code samples for open source programming languages and frameworks including Go, Ruby, Python, PHP, Java, Node.js and Sequelize, as well as Visual Studio Code and Jupyter notebook.

In terms of AI-driven automation capabilities, Db2 11.5 includes the new Augmented Data Explorer data exploration tool for data scientists, as well as natural language querying and the ability to deliver results in the form of data virtualizations and natural language summaries. Also new is machine-learning-based query optimization, while Db2 11.5 also provides integration with IBM's data virtualization functionality to enable access to data in multiple data platforms.

The first version of 11.5 includes only the first batch of functionality that is being delivered in relation to the synergy of data management and AI. There is more to come in terms of cost optimization and confidence-based query results, as well as exploiting Db2's graph, JSON and blockchain functionality, and integration with Apache Spark.

Licensing

The convergence of the underlying functionality means that Db2 is now available as a single download. By default, the download is the Db2 Community License, which is targeted at developers and can be used in development or production, either supported or unsupported, but is limited to four processor cores.

Customers can upgrade to the Standard License (which is targeted at departmental use cases and is limited to 16 cores and perpetual licensing), or the Advanced License (which is targeted at enterprise use cases and, therefore, has no restrictions on cores and is available with perpetual or subscription licensing. Each of the products based on Db2 (Db2, Db2 Warehouse, Db2 Hosted, Db2 on Cloud and Db2 Warehouse on Cloud), as well as those that use the Db2 Common SQL Engine (Db2 Big SQL, Db2 Event Store and the Integrated Analytics System) are available separately, while IBM also offers the combined Hybrid Data Management Platform.

The Hybrid Data Management Platform is designed to provide customers with an entitlement to all of the above products for a single price, enabling them to transition their usage as and when required, and is available with a perpetual license covering deployment on on-premises private cloud and traditional infrastructure, as well as a monthly subscription that covers deployment on public cloud, as well as on-premises private cloud and traditional infrastructure. IBM also offers bring your own license (BYOL) options for IBM Db2 on Cloud and IBM Db2 Warehouse on Cloud that provide cost savings for selected services for those with existing licenses.

Competition

IBM's primary competition for the combined Hybrid Data Management Platform comes from Oracle, Microsoft and SAP, while Teradata, Cloudera, Amazon Web Services, Google, Actian and Pivotal, along with a variety of smaller vendors, offer products and services that compete more directly with individual products in the portfolio.

Microsoft and Oracle must be considered the primary competition, and both have also been infusing their database offerings with AI-enabled functionality, as well as support for developing AI applications. While Microsoft offers Azure Cognitive Services as a complement to Azure SQL Database, Oracle has been actively positioning its core database functionality via Oracle Autonomous Database Cloud, which leans heavily on AI-based functionality combined with cloud infrastructure to deliver the 'autonomous' capabilities.

Even so, Microsoft and Oracle still provide distinct offerings for different use cases (Azure SQL Database, Azure SQL Data Warehouse and Azure Databricks, for example, or Oracle Autonomous Online Transaction Processing, Oracle Autonomous Data Warehouse and Oracle Big Data Cloud). As such, the Hybrid Data Management Platform offering is something of a differentiator in that it should enable users to easily adopt different underlying products as their needs change, rather than having to license multiple products or services.

But there's also a handful of other data warehouse vendors that are incorporating AI functionality either directly within the database or as a means for developing AI/ML-based applications. Teradata provides its Vantage offering and includes the ASTER analytics engine while Pivotal's Greenplum database includes the Apache MADlib open source analytics library. Still others include Micro Focus's Vertica with its embedded in-database ML functions and SAP HANA with its in-database analytical capabilities for running predictive analytics. Snowflake, the cloud-centric data warehouse, is similarly moving in the AI/ML direction but is taking the partnering route with tie-ins to Qubole and Databricks.

SWOT Analysis

STRENGTHS

IBM has a long-standing track record in data management and has expanded its database and data management portfolio over the years through a combination of internal development and acquisitions.

WEAKNESSES

Portfolio expansion had left the company with multiple brands and a fragmented code base, both of which have been addressed in recent years through the DB2 branding and Hybrid Data Management Platform launch.

OPPORTUNITIES

When AI becomes embedded within the data management system, it has the potential to improve database query accuracy and performance, for example, while the development of AI-based applications can be accelerated by support for data science tools.

THREATS

What exactly is an 'AI database'? IBM will have to make the case for its definition in the context of competing initiatives from the likes of Microsoft and Oracle, which are also expanding the automation they provide for database and query services.