

REPORT REPRINT

IBM unleashes updated DB2 with more powerful in-memory analytics

JASON STAMPER

21 JUL 2016

The company has announced the latest iteration of its database for Linux, Unix and Windows, which it says can handle both transactional and analytic workloads with aplomb.

THIS REPORT, LICENSED EXCLUSIVELY TO IBM, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, 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 REPOSTED, IN WHOLE OR IN PART, BY THE RECIPIENT, WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



©2016 451 Research, LLC | WWW.451RESEARCH.COM

IBM's recently released DB2 version 11.1 for Linux, Unix and Windows (LUW) is a hybrid database that IBM says can handle transactional and analytic workloads thanks to its BLU Acceleration technology, which features an in-memory column store for analytical workloads that can scale across a massively parallel cluster.

THE 451 TAKE

DB2 is a proven database for handling the most demanding transactional workloads. But the trend as of late is to enable relational databases to handle analytic queries more efficiently by adding an in-memory column store alongside to aggregate data and provide faster results. IBM's BLU Acceleration technology does exactly that. While BLU isn't brand new, the ability to spread the column store across a massively parallel processing (MPP) cluster of up to 1,000 nodes is a new addition to the technology. That, along with simpler monthly pricing options and integration with dashDB data warehousing in the cloud, makes DB2 for LUW, a very versatile database.

CONTEXT

The history of IBM's venerable DB2 database is both long and convoluted. It can be traced back to the early 1970s, when IBM researcher Edgar F. Codd published a paper on relational databases. The IBM San Jose Research center developed a relational database management system, System R, based on Codd's concepts in 1974. IBM also developed a relational database language called Structured English QUery Language or SEQUEL, which eventually morphed into Structured Query Language (SQL).

The name DB2, or IBM Database 2, was first given to the database in 1983 when IBM released DB2 on its MVS mainframe platform. IBM first made DB2 available on Intel and UNIX systems in 1993. Since then some of the acquired Informix technology was rolled into DB2, so that today it is technically an object-relational database. IBM has recently added a number of analytic capabilities into DB2 from its Netezza acquisition strengthening its in-database analytic capabilities.

Today IBM has different teams working on different platform editions of the DB2 database. The focus of this report is the LUW version, but there is also a version for z/OS mainframes, as well as IBM's i platform and VM/VSE. (The VM/VSE versions are no longer actively marketed, though they are still supported).

Each version is on a separate release cycle. DB2 for z/OS's latest release, for example, was version 12. That version offers improved in-memory techniques and compression-efficiency enhancements, making it capable of more than one million inserts per second with the ability to store up to 256 trillion rows in a single table. For the most demanding workloads, DB2 for z/OS is still IBM's recommended platform.

VERSION 11.1 FOR LUW

IBM says version 11.1 for LUW can handle mixed transactional and analytic workloads. This is thanks in part to the main object-relational data store that primarily handles transactional workloads, while a parallel in-memory columnar data store, known as BLU Acceleration, is optimized for analytical workloads. IBM told us that it has used some of the technology it garnered with its 2010 acquisition of analytics appliance vendor Netezza to improve the analytical capabilities of BLU Acceleration.

BLU Acceleration was available previously. However, IBM says it has enabled it to scale across 1,000 nodes with this release. (It couldn't previously be deployed across a MPP network cluster.) It says it has one early access customer that is planning to store 1.3 PB of data in the in-memory column store. IBM says that only frequently-accessed data is kept in memory, so that less frequently used data does not consume memory resources.

IBM pointed to an internal heavy analytics benchmark to demonstrate the performance of the BLU Acceleration in-memory technology. When running DB2 11.1 on an IBM Power E850 cluster, queries per hour were 1.92x faster when the cluster was increased from three to six nodes. In another test, queries per hour were about the same when the total data volume was doubled from 10TB to 20TB.

The company has also enhanced DB2's encryption capabilities with version 11.1. It has added support for the industry standard KMIP 1.1 for key management across the organization, but companies can still do local key management if they prefer.

IBM has also introduced a set of new licensing metrics for companies deploying DB2 in the cloud, or in hybrid cloud-on-premises deployments. DB2 Direct Standard Edition and DB2 Direct Advanced Edition offer a simplified license metric – the Virtual Processor Core (VPC) – that is sold as a monthly license charge. It can be bought online and delivered either in the cloud or on-premises. DB2 on Cloud can also integrate with IBM's cloud database and cloud data warehouse as a service, dashDB Transactional and dashDB Analytics.

IBM has also enhanced DB2 Connect with version 11.1. This connectivity software enables companies to integrate DB2 for LUW with DB2 for z/OS or DB2 for i. The latest version has improved developer productivity coupled with advances in performance, manageability and installation. Advanced editions of DB2 Connect also include the IBM pureQuery technology that helps to protect data from SQL injection attacks and other security threats.

Another new innovation is the DB2 Advanced Recovery Feature for LUW 11.1 – which is said to be a bundle of advanced database backup, recovery, and data extraction tools that can help to improve data availability, mitigate risk, and accelerate crucial administrative tasks. (New releases of each of the tools can still be bought separately). The bundle includes DB2 Merge Backup, DB2 Recovery Expert and the Optim High Performance Unload.

To make the most of the BLU Acceleration in-memory column store, enhancements have been made to the Data Server Manager tool to help organizations analyze their workloads and decide which tables would yield the greatest benefits when converted to BLU Acceleration column-oriented tables. A management reporting tool is also included that helps to track the percentage of user data that resides in BLU Acceleration column-oriented tables. BLU Acceleration is included in DB2 Advanced Enterprise Server Edition, DB2 Advanced Workgroup Server Edition (for smaller roll-outs), DB2 Direct Advanced Edition, and DB2 Developer Edition (for non-production use).

Those same editions also come with IBM's pureScale clustered-database technology that now offers support for high availability (HA) and disaster recovery (DR). This enables data to be mirrored from a primary pureScale database cluster to a second local or remote standby pureScale cluster synchronously, offering higher availability than the previous asynchronous clustering.

The pureScale capability also includes online fix pack update support. This means a company can apply DB2 11.1 fix packs to individual members running DB2 11.1 in a pureScale cluster without having to take the entire cluster offline, therefore minimizing any impact on users.

DB2 11.1 EDITIONS

There are eight different editions of DB2 11.1 for LUW. Enterprise Server Edition and Workgroup Server Edition, depending on the size of the task at hand, both have Advanced Editions that include the BLU Acceleration. The simplified virtual processor core (VPC) licensing editions are DB2 Direct Standard Edition and DB2 Direct Advanced Edition (with the BLU Acceleration).

Finally, there is a Developer Edition for non-production usage, and DB2 for Big Data. The latter includes DB2 as well as IBM's Hadoop-powered 'big data' platform, BigInsights. DB2 for Big Data is said to enable companies to move data between DB2 and BigInsights. There's data virtualization also, meaning relational and Hadoop-based data can be unified.

COMPETITION

DB2 for LUW's primary competition comes from Oracle and Microsoft. Oracle added a columnar in-memory feature with Oracle Database 12c In-memory Option. In contrast, Microsoft has had an in-memory capability since SQL Server 2014. Other popular competitors to DB2 include SAP's HANA and Sybase databases, as well as the popular open source relational databases, MySQL and PostgreSQL, both of which have commercial organizations offering support (Oracle and EnterpriseDB, respectively).

Since IBM is offering DB2 in new cloud-hosted and -managed SaaS options for hybrid environments, we'd expect companies to also consider Microsoft's Azure SQL Database, or Amazon Web Services with its Amazon Relational Database Service. Customers choosing that path can run Amazon Aurora, MySQL, MariaDB, Oracle, PostgreSQL or SQL Server databases in the Amazon cloud.

There are several database as a service (DBaaS) technologies based on MySQL, such as Rackspace Cloud Databases and Google Cloud SQL. We'd also list Tesora, Adapt and ScaleDB in the DBaaS space.

Intersystems' Caché object-relational database offers SQL support. Smaller companies claiming to handle transactional and analytical workloads include VoltDB, MemSQL, Altibase, Apache Geode (based on Pivotal GemFire), Clustrix, GridGain Systems, Deep Information Sciences, Hazelcast, Software AG, ScaleOut Software, JustOneDB and LeanXcale.

SWOT ANALYSIS

STRENGTHS

With proven reliability for transactional use cases, IBM has stepped up DB2's analytic capabilities even further with cluster-capable in-memory column stores.

WEAKNESSES

The different versions of DB2 for (LUW), z/OS, i and VM/VSE are not in lockstep, which may be less than ideal for companies running more than one hardware platform.

OPPORTUNITIES

We see a growing interest in companies wishing to run mixed transactional and analytic workloads as a way of reducing complexity and speeding analytic insight.

THREATS

Both Oracle and Microsoft have new versions of their databases out this year. New deals in the relational database market remain highly competitive. Open source remains an option for customers and an ever-present threat to established company products and solutions.