NoSQL keeps rising, but relational databases still dominate big data

by **Matt Asay** in **Big Data** on April 5, 2016, 5:53 AM PST

Even though MongoDB and Cassandra keep winning converts, enterprises are keeping their RDBMSes around, and will do so for quite some time.

Image: iStockphoto/Sergey Nivens

NoSQL promised to upend the database market as big data forced a sea change in how we think about and manage data. Several years into the big data revolution, that promise remains unfulfilled.

Oh, sure, NoSQL is having an impact on the $46 billion database market--still just 3% of the market, but growing at a rapid pace even as more traditional relational databases inch up by 5.4%, according to IDC (http://www.valuewalk.com/2014/07/oracle-corporation-orcl-launches-big-data-sql/). But, by that same measure, as well as the updated DB-Engines database popularity rankings (http://db-engines.com/en/ranking), relational databases still dominate big data.

SEE: Research: Big data and IOT - Benefits, drawbacks, usage trends (http://www.techproresearch.com/downloads/research-big-data-and-iot-benefits-drawback-usage-trends/) (Tech Pro Research)

On current trends, then, we can expect NoSQL and relational databases to share the big data winner's podium for many years to come.

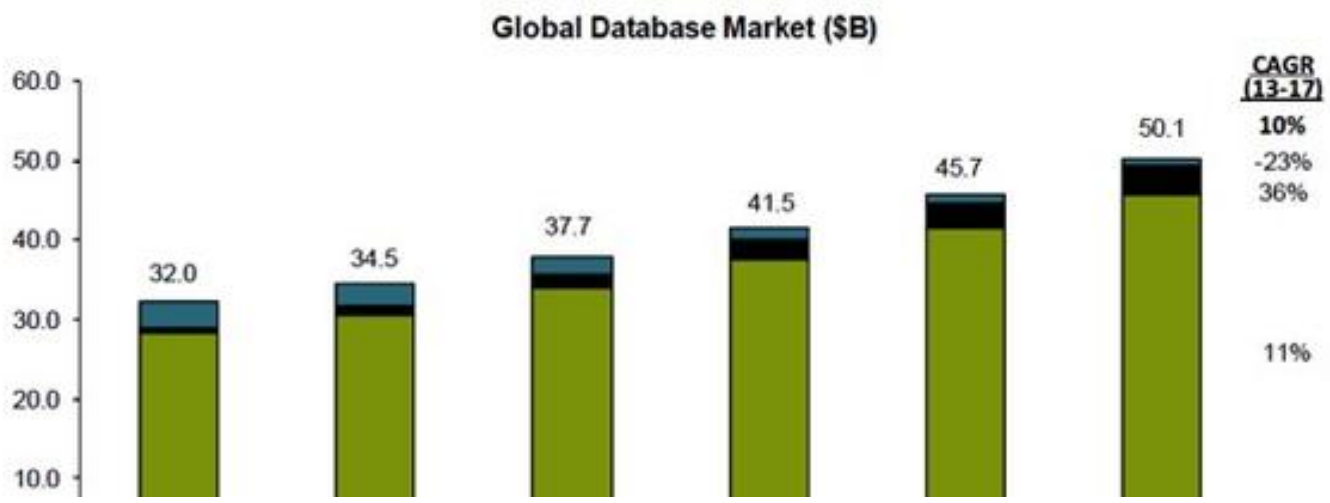
NoSQL rising

NoSQL is a response to the dramatic growth of unstructured data within the enterprise, and promises to be the biggest challenge to the established relational database management system (RDBMS) hegemony. Though Oracle currently controls 40% of the paid database market, analyst Curt Monash warns (http://www.dbms2.com/2015/12/31/oracle-as-the-new-ibm-has-a-long-decline-started/) that NoSQL is one of the biggest threats to its crown:

"[T]here are basically three things that can seriously threaten Oracle's market position, [the first of which is] growth in apps of the sort for which Oracle's RDBMS is not well-suited. Much of 'Big Data' fits that description."

Even so, there's a long way for NoSQL to go, as IDC data highlights:

Exhibit 5
Hadoop and NoSQL accounted for only ~3% of total market, but are expected to grow rapidly going forward.



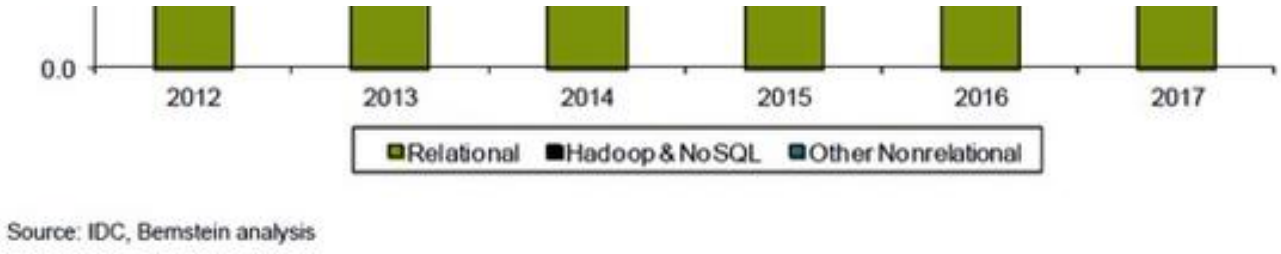


Image: IDC

Despite the uphill battle, DB-Engines' broad view into database popularity (http://db-engines.com/en/ranking) shows a continued narrowing of the gap by the top NoSQL databases, including MongoDB, Datastax-sponsored Apache Cassandra, and Redis:

303 systems in ranking, April 2016

Rank			DBMS	Database Model	Score		
Apr 2016	Mar 2016	Apr 2015			Apr 2016	Mar 2016	Apr 2015
1.	1.	1.	Oracle	Relational DBMS	1467.53	-4.48	+21.40
2.	2.	2.	MySQL	Relational DBMS	1370.11	+22.39	+85.53
3.	3.	3.	Microsoft SQL Server	Relational DBMS	1135.05	-1.45	-14.07
4.	4.	4.	MongoDB	Document store	312.44	+7.11	+33.85
5.	5.	5.	PostgreSQL	Relational DBMS	303.73	+4.10	+35.41
6.	6.	6.	DB2	Relational DBMS	184.08	-3.85	-13.56
7.	7.	7.	Microsoft Access	Relational DBMS	131.97	-3.06	-10.22
8.	8.	8.	Cassandra	Wide column store	129.67	-0.66	+24.78
9.	9.	↑ 10.	Redis	Key-value store	111.24	+5.02	+16.69
10.	10.	↓ 9.	SQLite	Relational DBMS	107.96	+2.19	+5.67

Image: DB-Engines

This growth is happening against the backdrop of a slowdown in the RDBMS market. Oracle, which sits atop the database popularity and revenue rankings, keeps deriving less and less of its revenue from new license sales, as Redmonk analyst Stephen O'Grady captures (https://www.techrepublic.com/article/two-new-ceos-dont-add-up-to-meaningful-change-at-oracle/):

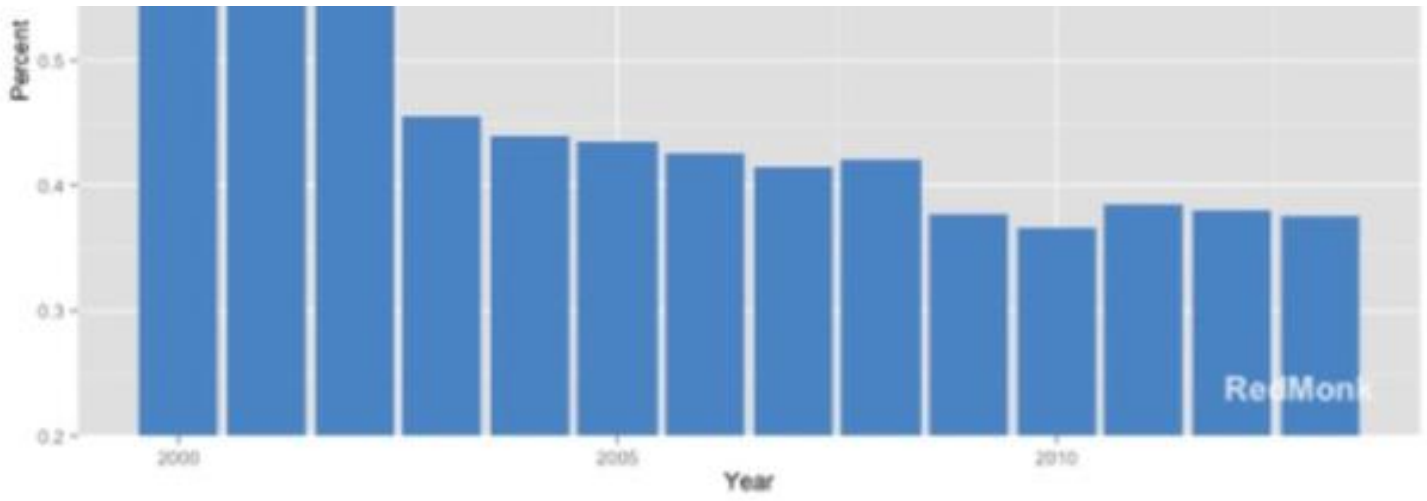


Image: Redmonk

However, the other side of the market leaders' ledger shows huge growth for MySQL, the open source RDBMS Oracle acquired with the purchase of Sun Microsystems (open source database growth has clocked 31%, according to IDC). While Oracle makes far less money from MySQL subscription sales than it does its eponymous database, MySQL is arguably doing more to keep it relevant to modern workloads (and the developers who love them).

RDBMS keeping pace

For all of NoSQL's rise, however, RDBMS isn't exactly falling. There are several reasons why this is the case.

The first is that, as much as enterprises may want to embrace NoSQL to tame mounting quantities of unstructured data, most of their workloads remain transactional in nature, which is the sweet spot for RDBMS.

More about Big Data

- Data management: A cheat sheet (https://www.techrepublic.com/article/data-management-a-cheat-sheet/)
- How to integrate robotic process automation in big data projects (https://www.techrepublic.com/article/how-to-integrate-robotic-process-automation-in-big-data-projects/)
- Programmer Hadley Wickham touts diversity of R community (https://www.techrepublic.com/article/programmer-hadley-wickham-touts-diversity-of-r-community/)
- How to choose the right data analytics tools: 5 steps (https://www.techrepublic.com/article/how-to-choose-the-right-data-analytics-tools-5-steps/)

Another reason is that analytics tooling for NoSQL is still in its infancy. As Gartner analyst Lynn Robison points out (http://www.informationweek.com/big-data/big-data-and-rdbms-can-they-coexist/d/d-

id/1324939?_mc=sm_iwk&hootPostID=0a32c68b759c64888ffc0887aaab6cf7), NoSQL-friendly analytics tools are not user-friendly analytics tools, and "It will take years for analytical tools to mature and become accessible to people who are not in data science."

A third reason is the hardest to quantify and to overcome: Culture. Enterprises have spent the last 30-plus years living with relational databases. It's hard to change that overnight.

Finally, a fourth reason: It turns out an RDBMS is sometimes the best solution for a particular problem. Just ask Facebook's data chief, Ken Rudin (http://searchcio.techtarget.com/opinion/Relational-databases-are-far-from-dead-just-ask-Facebook):

"If we look at the granularity of the data, we keep the lowest level of grain in our Hadoop system. So whenever you want to look at something at the lowest level of detail, Hadoop is optimized for that. However, if we want to look at transformed data and aggregated data, relational is easier for doing that."

So, real-time monitoring is done in Hadoop and associated NoSQL databases, but for longer-term trending analytics, that's RDBMS.

SEE: Big data privacy must be fixed before the revolution can begin (http://www.zdnet.com/article/big-data-privacy-must-be-fixed-before-the-revolution-can-begin/) (ZDNet)

Two 'complete game changers'

Professor Michael Franklin, one of the industry's foremost database experts, has hailed (https://www.techrepublic.com/article/nosql-is-a-complete-game-changer-declares-database-expert/) NoSQL as a "complete game-changer," largely because of the flexibility of its schema. But, that very flexibility also leaves the door wide open for a relational database, which can perform fast, powerful queries against data that has been neatly packaged into rows and columns.

In other words, RDBMS and NoSQL are both "complete game changers," and together they can support winning big data strategies.

Also see