



RESEARCH NOTE

Oracle cranks up the heat in the MySQL market

Bringing high-performance analytics to MySQL

Executive Summary

Trigger

Oracle owns MySQL, but paradoxically, has never commanded much mindshare in the MySQL cloud space. But with a new service, it is coming out swinging. Oracle has launched a new fully managed MySQL service on Oracle Cloud Infrastructure (OCI) that is not simply a carbon copy of the many MySQL cloud services already available from each of the incumbent cloud providers. At the core is “HeatWave,” a new in-memory analytics accelerator developed for the MySQL Database Service that takes MySQL into new territory by integrating analytics into a database traditionally known for transaction processing.

Our Take

The MySQL landscape needed a shakeup. Until now, it was considered the default go-to open source database for online transaction processing (OLTP) for the scenarios not requiring the more sophisticated capabilities of “enterprise” databases such as Oracle or PostgreSQL. And, owing to lack of features such as materialized views, MySQL has not been seriously considered for analytics/data warehousing use cases. However, given the popularity of MySQL with its large skills base, there was a vacuum waiting to be filled. Oracle has been actively developing MySQL for over a decade, but in the MySQL cloud ecosystem, Oracle was overshadowed by cloud providers offering their own MySQL DBaaS services. To enter a MySQL DBaaS market that was already pretty crowded, Oracle would have to sharply differentiate itself. Oracle’s new “HeatWave” engine does just that, turning up the heat on rival cloud MySQL services by not only adding analytics to the mix, but incorporating hardware and software optimizations that significantly accelerate performance. With the new service, MySQL customers can use the same familiar database for their OLTP and OLAP workloads, while eliminating the cost, complexity, and risk of the ETL processes that are required when using separate databases. Aggressively priced, Oracle MySQL Database Service customers should get a good bang for their buck.

A low-key stakeholder

MySQL was first conceived in the 1990s as a relatively simple, easy-to-implement relational database for handling transactions for web applications. It became the data pillar of the open source LAMP Stack (which included Linux, Apache webserver, MySQL, and either Perl, Python, or PHP programming languages).

To this day, MySQL remains an incredibly popular database. According to the db-Engines December 2020 ranking, MySQL is outranked in popularity only by the Oracle Database. Although the database is over 25 years old, in 2019, db-Engines ranked MySQL as the leading gainer in popularity during that year.

So, according to the rankings, Oracle owns the top databases for the entry-level and enterprise ends of the market. While Oracle has a clear commercial presence with its flagship database, offered in flavors ranging from Autonomous Data Warehouse to Exadata Cloud Service, in the MySQL cloud ecosystem, it has kept a fairly low profile until now. MySQL came to Oracle via the 2009 Sun Microsystems acquisition and Oracle continues to lead its development. But, up until now, if you look at commercial activity in the cloud, offerings from incumbent cloud platform providers have dominated the market. Microsoft Azure, Google Cloud, and Alibaba Cloud each offer one managed MySQL service, while AWS offers two. Third parties, such as Percona, are also active in offering their own supported MySQL distributions. There are parallels with Oracle's early days, but with the roles reversed; there, IBM invented the relational database, but Oracle was the first company to exploit it commercially.

With the new Oracle MySQL Database Service and its integrated HeatWave analytics engine, Oracle is not only aiming to change the script, but also completely redefining what MySQL is.

Expanding MySQL's horizons

Oracle's new MySQL service starts with the same InnoDB storage engine that has been a pillar of most MySQL implementations. In spite of the fact that MySQL supports plugging in alternative storage engines, InnoDB has remained the most prevalent. However, InnoDB is not optimized for analytics; furthermore, the MySQL platform lacks support of materialized views that are critical for many analytics databases (they speed access to common data splits). For those reasons, Oracle needed to add a separate storage engine operating in parallel to handle analytics while maintaining compatibility with the transaction engine.

It does so by adding a new, in-memory analytics accelerator developed for MySQL Database Service: HeatWave. It's a distributed, scalable, in-memory, hybrid columnar, query-processing engine designed for fast execution of analytic queries. Customers might be forgiven for associating this with the columnar technology underpinning Oracle Database In-Memory that debuted with 12c. However, HeatWave is a completely different in-memory hybrid columnar implementation, designed for a completely different database platform.

With HeatWave, data is persisted in the MySQL InnoDB storage engine. Any updates to the tables are automatically propagated to the memory of the HeatWave nodes in real-time. This allows subsequent queries to always have access to the latest data. This is done behind the scenes by a lightweight change propagation algorithm that keeps pace with MySQL data update rates; the process is more efficient than data replication. Additionally, there's no need to index the data before running analytics queries.

The size of the HeatWave cluster depends on tables and columns required to load, and the compression achieved in memory for this data. A node count estimate is generated using Machine Learning, initially or at any time after to adjust the number of nodes.

There are several other tricks for analytics up Oracle's sleeve with the new HeatWave analytics engine. The Oracle service micro-partitions the data in the hybrid column store to allow massive query parallelization. Additionally, it introduces vector processing, a capability which allows a sequence of instructions to be executed on a group of rows. With vector processing, the overhead of execution is spread over multiple rows, which accelerates performance. There are relatively few other commercial platforms using vector processing (Actian is one of them). The combination of conventional in-memory columnar, micro partitioning, and vector processing makes HeatWave unique.

There are further optimizations made for the underlying Oracle Cloud Infrastructure (OCI) on which MySQL Cloud Service runs. Compute and communication between compute nodes are optimized for network bandwidth, while sizes of partitions are optimized to fit the available cache. There are additional optimizations for internal hardware instructions. These help HeatWave obtain very good scalability across a large number of servers.

Once users submit a query to the MySQL database, the MySQL query optimizer decides if the query should be offloaded to the HeatWave cluster for accelerated execution. HeatWave accelerates performance of MySQL by 400X for analytics queries, and scales out to thousands of cores. Claiming superior performance, Oracle has run its own TPC-H benchmark comparing MySQL Database Service with HeatWave vs. Amazon Redshift. This is not an opaque benchmark; Oracle encourages customers to run it for themselves, making all the scripts publicly available on GitHub at the following link:

<https://github.com/oracle/heatwave-tpch>

Oracle is not the first to extend open source

The issue of open source purity is debated frequently, but for most enterprises, the real question is over compatibility: will their programs run on the extended platform without being broken, or will they need different skill sets because there are differences with the so-called pure open source platform. For most commercial open source offerings, these questions are academic – and that is especially true for Oracle's extension of MySQL. The real difference is that extended platforms may support greater scale or functions.

In the MySQL world, Oracle is not the first to extend the platform. However, it does so in a different way compared to Amazon Web Services (AWS), which offers a premium service for high-scale OLTP with Aurora. AWS modifies MySQL by using a proprietary storage engine, maintaining compatibility at the API level with MySQL 5.6 and MySQL 5.7.

There are even more precedents in the PostgreSQL world where there are numerous analytic databases (e.g., Amazon Redshift, Pivotal Greenplum, IBM Netezza) and some transaction databases (Azure Database for PostgreSQL Hyperscale, which incorporates extensions from the Citus DB acquisition) that trace their roots back to PostgreSQL. Each of these implementations are forks from the original PostgreSQL open source project.

HeatWave takes advantage of MySQL's pluggable storage engine architecture; it plugs in just like any other MySQL storage engine and is natively integrated with InnoDB, MySQL's default storage engine. It is implemented in a way that shields all the low-level implementation details at the storage level from the end users. As a result, users can manage both storage engines in the Oracle MySQL Database Service with the same management tools including the Oracle Cloud Infrastructure console, REST API, and command line interface. Users connect to HeatWave via MySQL, and as a result, all standard tools and standard-based ODBC/JDBC connectors work without modification. HeatWave supports the same ANSI SQL standard and ACID properties as MySQL and supports diverse data types, and with no changes to application code. HeatWave supports the same BI and data visualization tools as MySQL Database, such as Oracle Analytics Cloud, Tableau, and Looker.

The bottom line for Oracle MySQL customers is that, if they are migrating from an existing MySQL instance, their programs will work. The only difference that, in the Oracle service, they will also be able to run complex analytic queries.

Takeaways

It's ironic that until now, the company that has owned the MySQL database and led its development was among the least known for it in the cloud. It's about time for Oracle to change the narrative.

It is not simply adding just another me-too MySQL cloud service to an already-crowded market. It is differentiating it, both in features and price. Oracle MySQL Database Service, with its integrated HeatWave engine, is the only MySQL cloud service that supports OLTP and OLAP workloads. The big benefit is with transaction and analytic workloads relying on the same data. Because data stays within the same database, it dispenses with the need for ETL processes. HeatWave is a versatile engine for analytics, enabling in-memory vectorized processing, massive inter-and intra-node parallelism, and distributed query processing; the engine chooses the optimal path for the query. Because it is optimized for OCI, Oracle can further speed MySQL Cloud Service performance.

The new service expands MySQL into new analytic territory, and it is doing so in a way that won't break applications for existing MySQL customers. Sealing the deal, Oracle is aggressively pricing the new service at a level that it claims is one-third the cost of Amazon

RDS for MySQL and Azure Database for MySQL. We have not independently verified pricing, as it depends on variables such as whether instances are on-demand or reserved, etc. However, we certainly believe that Oracle will not only match but also actively seek to underprice its rivals. This is consistent with Oracle's pricing strategy for Autonomous Database, where it goes as far as offering savings guarantees versus rivals; it can do so because it controls the running of the database.

The Oracle MySQL Database Service, with its high-performance HeatWave analytics engine, is a well-thought out debut cloud offering that turns up the heat on rival MySQL services .

Author

Tony Baer, Principal, dbInsight

tony@dbinsight.io

Twitter @TonyBaer

About dbInsight

dbInsight LLC provides an independent view on the database and analytics technology ecosystem. dbInsight publishes independent research, and from our research, distills insights to help data and analytics technology providers understand their competitive positioning and sharpen their message.

Tony Baer, the founder and principal of dbInsight, is a recognized industry expert on data-driven transformation. *Analytics Insight* named him one of the [2019 Top 100 Artificial Intelligence and Big Data Influencers](#). His combined expertise in both legacy database technologies and emerging cloud and analytics technologies shapes how technology providers go to market in an industry undergoing significant transformation. His regular ZDnet "Big on Data" posts are read 25,000 – 30,000 times monthly.