Publication date:
24 Aug 2021
Author:
Bradley Shimmin, Chief Analyst, AI & Data Analytics

Oracle puts MySQL on autopilot with new AI/ML-fueled capabilities for both transactional and analytical workloads

Brought to you by Informa Tech

Table of Contents :

Omdia view

Summary

Oracle's recently introduced Oracle MySQL Autopilot as an element within its in-memory query optimizer, MySQL HeatWave, adds an important twist to Oracle's already distinctive take on cloud-native database architectures. MySQL HeatWave extends the company's focus on automation to its fully managed MySQL Database Service running on Oracle Cloud Infrastructure (OCI).

Blending database workloads at scale

Oracle has been on a crusade over the past few years, trumpeting the importance of autonomous computing, meaning the aggressive application of artificial intelligence (AI) and machine learning (ML) to the multitudinous problems of deploying, operating, managing, and maintaining an enterprise-scale database platform. To date the company has made its case with services such as Oracle Autonomous Database, proving internally that the application of AI to problems like applying security patches in a timely and reliable manner or optimizing SQL queries based upon current and predicted database workloads can lead to significant reduction in management costs.

More than automation, with solutions like Oracle Autonomous Database, the company has also been championing the idea of multimodal databases, which enable a single database instance to support disparate workloads, be those data types (structured, semi-structured, or unstructured) or workloads (operational, analytical). Why combine data types and workloads within a single database? Historically, database technologies have demanded specialization. Companies wishing to do online transactional processing (OLTP) at scale would never try to accommodate online analytical processing (OLAP) as doing so would surely degrade OLTP performance. Similarly, companies looking to tackle graph analysis at scale would never try to force a traditional relational database to accommodate the exponential demands of analyzing huge numbers of nodes and edges.

Unfortunately, this perception of specialization over generalization has created many challenges for enterprise data practitioners, namely a proliferation of data silos, which not only adds complexity but also the cost in both time and money -- the time it takes to move operational data to a data warehouse and cost in the expense of moving that data, not to mention the associated security risks and analysis of stale data. The industry wide rush to embrace the cloud has of course greatly exacerbated this situation. Data ingress and egress charges between databases and cloud platforms themselves can dwarf the comparative cost of managing a single database.

To its credit, Oracle has been able to show customers the value in unifying disparate data types and data workloads within its flagship database running on OCI, and more recently, even on premises with Oracle Exadata Cloud@Customer and Dedicated Region Cloud@Customer. But before the introduction of Oracle MySQL HeatWave late last year, the company had not extended this same benefit to MySQL, the company's hugely popular transactional database included with the Sun Microsystems acquisition in 2009. MySQL HeatWave, which is an in-memory query accelerator, sits on top of MySQL Database, enabling the OLTP database to handle complex, analytical queries that would otherwise fall to a secondary, analytics-focused database. In short, MySQL HeatWave does a lot more than speed up queries by scaling to thousands of cores. It allows users to query operational data in real time without first forcing them to extract, transform, and load (ETL) transactional data into a separate database and without requiring any sort of application code changes.

Bringing the autonomous database to the enterprise

MySQL HeatWave was already equipped with AI/ML capabilities geared toward performance and scale, a point Oracle is eager to leverage as a means of competing with several cloud-native databases including Google BigQuery, Snowflake, Amazon Aurora and Redshift, Microsoft Azure Synapse, and other MySQL-based databases. Now, with the introduction of MySQL Autopilot for MySQL HeatWave, Oracle is bringing in AI/ML-driven automation and optimization capabilities to bear across a number of database management challenges including provisioning, data ingestion, query execution, and error handling.

Generally available now free of charge to all MySQL HeatWave customers, MySQL Autopilot does not just use static ML models to optimize the MySQL HeatWave query engine. Rather, as an exclusive capability on OCI, it uses Oracle AutoML technology to model and re-model numerous aspects of the database such as memory usage, network load, and execution times as those metrics change over time. This makes MySQL Autopilot a true learning machine that performs its core tasks with increasing accuracy the longer it is in use. In support of provisioning, for example, it can predict the number of MySQL HeatWave nodes that are required to run a given workload by sampling the table data adaptively over time. This basically frees customers from having to manually calculate optimal cluster sizes. The database can just adapt to demand automatically.

MySQL Autopilot automates several capabilities beyond auto provisioning, including:

- **Parallel loading**: Predict the optimal degree of parallel processing to optimize load time and memory utilization for each table loaded into MySQL HeatWave.
- **Data placement**: Optimize query performance by predicting the specific table column that should be partitioned in-memory and also predict the expected improvement with this change.
- **Encoding**: Minimize cluster size and improve query performance by automatically determining the best representation of columns loaded into MySQL HeatWave.
- **Scheduling**: Determine which queued queries are short- or long-running, prioritizing short-running queries in order to minimize the wait times for the shorter running (interactive) queries.
- **Change propagation**: Determine the most optimal time to propagate MySQL Database changes to the recently introduced HeatWave scale-out storage. Note that changes from MySQL are always propagated to the HeatWave memory in real-time.
- **Query time estimation**: Ascertain query execution time prior to running a given query, even for the first time.
- **Query plan improvement**: By analyzing query execution metrics over time, improve the execution plan for queries incrementally over time.
- **Error handling**: Automatically provision new nodes and reload data as necessary should one or more MySQL HeatWave nodes become unresponsive due to a hardware failure.

If there is one theme running throughout these numerous AI/ML-based features, it's optimization. More specifically, for OCI customers running MySQL Database Service, the combination of HeatWave and Autopilot technology promises to do two things. First, it will power up their operational databases, and in so doing, it enables them to start running analytical workloads right next to their traditional analytical workloads without having to take performance hits or incur additional costs to overcome performance limitations. That marks a major win for existing MySQL Database Service customers, especially those already embracing cloud-native computing on top of OCI. Rather than having to perform ETL and wait for operational data to load into a separate data warehouse in order to analyze and respond to changes in the data as they are written to the operational database, users can simply query that data in place and in real-time via MySQL HeatWave's in-memory query engine as optimized by MySQL Autopilot.

Enterprise buyer considerations

From a go-to-market perspective, Oracle's intent with MySQL HeatWave and MySQL Autopilot is quite clear. Customers invested in any MySQL derivative, whether on premises or in cloud, should move to or stick with MySQL Database Service on OCI. Assuming the cloud is an option, there are few downsides, as Oracle can promise to enhance security, reduce management, and lower operational costs while improving performance -- all while expanding traditional OLTP workloads to include complex analytics queries. Even for customers who need to keep their data on-premises, the Oracle MySQL Database Service provides a hybrid deployment setup where customers' on-premises data can be replicated to OCI using standard MySQL replication. Customers can then run MySQL HeatWave on the replicated instance in the cloud.

MySQL HeatWave and MySQL Autopilot can support any existing MySQL implementation without forcing users to rewrite their applications or queries, which creates an open invitation to customers currently invested in a MySQL-compatible database such as Amazon Aurora or MariaDB on non-OCI cloud platforms to do their own benchmarking to see if Oracle's automated approach can save them money. Second, Oracle is promoting the enticing notion that customers currently maintaining two databases -- MySQL for transactions and a separate analytics database such as Snowflake, Amazon Redshift, or PostgreSQL -- in support of a given use case, can throw out their second database and all of the overhead, including ETL processes, that goes along with it.

While multi-cloud integration, support, and management loom on the horizon, as evidenced by early efforts such as MongoDB Atlas, shift in allegiance from one database on one cloud platform to another remains a significantly non-trivial endeavor. For that reason, Oracle will have to not only work on its support for multi-cloud computing but also demonstrate the value proposition of the numerous AI-fueled innovations that make up OCI. Ultimately the question for users revolves around their own unique centers of gravity when it comes to their commitment to and investment in both database and cloud platform provider. MySQL customers just moving to the cloud will find this a straightforward proposition with OCI coming out the clear winner.

Those dedicated to on-premises MySQL can turn on MySQL replication to move data from on-premises to the cloud and run MySQL HeatWave. But for those already steeped in rival cloud platforms, particularly Google Cloud Platform (GCP), Microsoft Azure, and AWS, the decision to move to OCI will likely come down to whether or not Oracle's performance improvements and cost reduction claims stand up to their own investigations. To that end, Oracle has posted the third party benchmarks included as part of the MySQL HeatWave announcement on GitHub. Potential customers can use this information to replicate Oracle's benchmarks.

Open source stewardship

There is also the not-insignificant matter of the company's stewardship of the open source MySQL project. Oracle must continue its investment in MySQL itself, forging points of synergy, for example, between MySQL Community Edition and emerging functionality such as HeatWave to prove to the industry that it remains the best steward of this important open source community. For Oracle this stewardship naturally revolves around ecosystem participation. The MySQL engineering team makes quarterly contributions to the open source community, ensuring that the general public has the core advances to the database itself (note, however, that HeatWave runs only on OCI and is not part of the open source contributions). Likewise, Oracle ensures that enterprise customers have access to the most recent MySQL open source releases. For instance, MySQL Database Service with HeatWave runs the latest MySQL version (currently version 8.x).

For many enterprise buyers, what matters most with open source stewardship is not code contributions but rather choice. From Oracle's perspective this means ensuring that with MySQL Database works just like Oracle Linux and Oracle Database: if customers want MySQL on-premises, including the free open source version, the standard code is actually the same as that used elsewhere. The core engine can be run on-premises, in a Docker container, or on a VM or on a laptop— all of these use the same code that runs within MySQL Database Service. It is this kind of inherent flexibility that drives new capabilities like MySQL HeatWave and MySQL Autopilot and allows customers to put those new capabilities to work across a broad spectrum of use cases and requirements without having to worry that future market trends may lead them away from their existing investments.

Appendix

Author

Bradley Shimmin, Chief Analyst, AI & Data Analytics

askananalyst@omdia.com

Citation policy

Request external citation and usage of Omdia research and data via citations@omdia.com.

Omdia consulting

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Omdia's consulting team may be able to help you. For more information about Omdia's consulting capabilities, please contact us directly at consulting@omdia.com.

Copyright notice and disclaimer

CONTACT US

omdia.com

askananalyst@omdia.com