FUTURUM.

MySQL Database Service with HeatWave: A Massively-scalable Integrated Query Accelerator on Oracle Cloud Infrastructure

Why MySQL HeatWave running on Oracle Cloud Infrastructure is the best value bar none for MySQL Database applications

Q3 2021

DANIEL NEWMAN
Founding Partner + Principal Analyst

SHELLY KRAMER
Founding Partner + Lead Analyst

RON WESTFALL
Senior Analyst + Research Director

Published: August 2021

IN PARTNERSHIP WITH

ORACLE

Table of Contents

Introduction

Oracle Unveils Innovations in MySQL HeatWave to Bring the Best Price/ Performance Proposition to Market

MySQL is the most popular open-source SQL database used in the world and is widely used across the world's leading Web applications such as Uber and Facebook. MySQL's open-source relational database (DB) management system design enables multiple users to manage and create DBs according to their rapidly evolving business and technical requirements.

As a result, we see the Oracle MySQL Database Service with Heatwave playing a pivotal role in driving innovation across the MySQL ecosystem, consistently fulfilling the most demanding and complex query tasks. Specifically, Oracle's MySQL HeatWave offering is an in-memory query accelerator for MySQL Database Service that enables customers to use a single, high-performance MySQL database for both transactional and analytical workloads. It assures the smooth running of complex queries against a customer's MySQL DB without any ETL (Extract, Transfer, Load) operations requirement. Through HeatWave's ability to scale to thousands of cores, customers can maintain a single MySQL DB for all their applications with the built-in assurance that they will all work without requiring any changes.

The new MySQL HeatWave release, including MySQL Autopilot ML-powered automation capabilities such as auto parallel loading, auto scheduling, and auto provisioning, scale out data management, and a cloud-first design, is the best-suited industry offering to meet the topmost concerns and demands of MySQL customers globally. In examining the new MySQL Database Service with HeatWave solution, we analyze the following elements:

- MySQL HeatWave's cloud-first design
- Machine learning's key role in driving MySQL HeatWave design
- MySQL Autopilot: ML-fueled Automation of MySQL Database Services
- Scale-Out Data Management for MySQL HeatWave
- Price/Performance Competitive Edge
- Why Window Function Support is Key
- Turbo-boosting both OLTP (Online Transactional Processing) and OLAP (Online Analytical Processing) Workloads

Executive Summary

- The MySQL Heatwave solution delivers the extensive automation, unparalleled performance, scalability, lower costs, blended workloads, and security that MySQL customers are prioritizing.
- Oracle ML models enable MySQL HeatWave to adapt to changing workloads without the need to reprogram the rules and can make recommendations based on the state of the system in the future, including predictions for improvement with the recommended change as well as exercising the ability to provide an explanation for the recommended change.
- The MySQL Autopilot architecture, available exclusively on OCI, introduces innovations that build-in ML capabilities to ensure MySQL HeatWave performance breakthroughs.
- The capabilities integral to MySQL HeatWave allow Oracle to demonstrate the most compelling price/performance advantages against its main MySQL rivals that we've seen to date.
- MySQL HeatWave's Scale-Out Data Management supports the reloading of any amount of data in constant time.
- TPC-H, TPC-DS, and the new fully transparent CH-benCHmark benchmarks demonstrate HeatWave's performance, price and scale advantages over all other MySQL and cloud databases. They are clear, repeatable tests that provide the hardcore facts which support Oracle's claims.

MySQL Database Service with Heatwave Overview: Purpose-designed for the Cloud

An overview of Oracle's MySQL DB Service with Heatwave proposition we believe is warranted to fully understand why the solution delivers such clear-cut price and performance advantages. For starters the MySQL HeatWave solution is designed from its inception for the cloud with many innovations around new distributed algorithms for query processing. It has been optimized specifically for OCI (Oracle Cloud Infrastructure), enabling the attainment of breakthrough capabilities such as scale-out processing and the ability to use commodity cloud services that offer the least expensive compute, storage, networking, and Virtual Machines (VMs). In addition, the solution makes pervasive use of machine learning (ML) to assure such price/performance outcomes.

We see MySQL customers putting topmost selection emphasis on automation throughout the entirety of their MySQL deployment, capitalizing on performance advances required to manage massive data growth, and using HeatWave capabilities to accelerate more queries. Moreover, they are requiring support for larger data sizes, an improvement for mixed OLTP and OLAP workload performance, and the encryption of their MySQL network communications.

Through a cloud-first development strategy, we believe the MySQL Heatwave solution delivers the extensive automation, unparalleled performance, impressive scalability, lower costs, extensive automation, blended workloads, and security that MySQL customers are prioritizing.

MySQL Database Service with Heatwave Overview: Machine Learning's Vital Role

Oracle's strategic commitment to machine learning (ML) centers on Oracle Machine Learning technology and AutoML. Through Oracle Machine Learning, Oracle moves the algorithms to the data and runs ML within the DB where data resides. We view Oracle's approach as minimizing or eliminating data movement, advancing scalability, preserving data security, and accelerating time-to-market model deployment.

In contrast to rule-based models, HeatWave's Autopilot models are specific to the instance, and not static for all the users. Oracle's Autopilot models can adapt to changing workloads without the need to reprogram the rules and can predict the amount of improvement expected with the predicted change. By providing an explanation behind the change recommendation, we see Oracle further distinguishing the integral role its ML technology plays in differentiating the MySQL HeatWave solution.

Oracle's Autopilot models can adapt to changing workloads without the need to reprogram the rules and can predict the amount of improvement expected with the predicted change.

MySQL Autopilot: ML-fueled Automation of MySQL Services

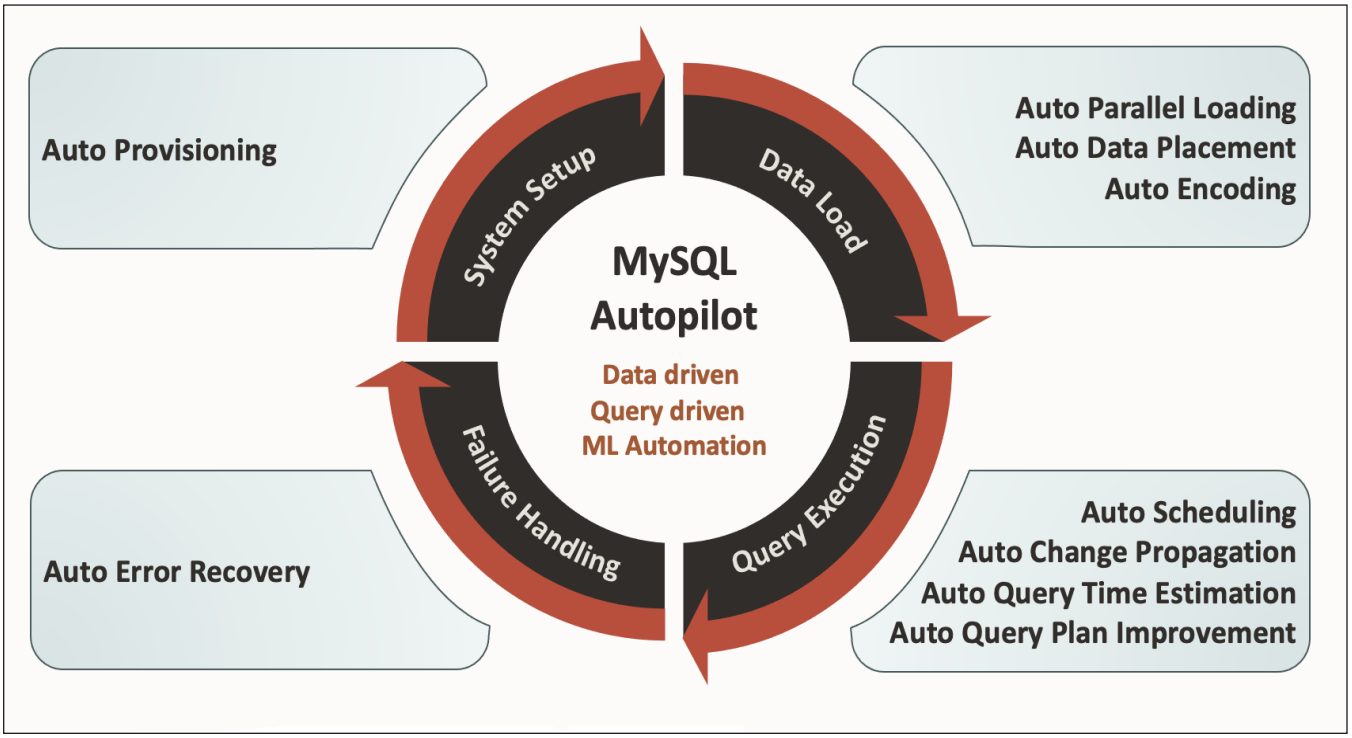
The MySQL Autopilot architecture, available exclusively on OCI, introduces innovations that build-in ML capabilities required to ensure MySQL HeatWave breakthroughs. This includes integrating ML-based decision logic throughout entire service lifecycles, AutoML models to build and train models, statistics extrapolation, and adaptive sampling across components, such as InnoDB, HeatWave, MySQL Optimizer, and Query Execution logs, that feed into statistics extraction.

Key MySQL Autopilot features include:

- **Auto Provisioning:** Supports system setup.
- **Auto Parallel Loading, Auto Data Placement, Auto Encoding:** Improve data load performance.
- **Auto Error-Recovery:** Provides immediate failure handling.
- **Auto Scheduling, Auto Change Propagation, Auto Query Time Estimation, Auto Query Plan Improvement:** Optimizes query execution.

MySQL Autopilot Features

Machine learning based automation



Specifically, auto provisioning delivers ML-driven prediction of memory usage to estimate cluster size, enabling customers to move away from analyzing DB schemas and tables to guess the cluster size needed. The auto data placement capabilities use ML prediction to attain optimal in-memory partition columns. As such, the system can predict optimal columns to partition data in-memory based on recent queries, reducing or eliminating data movement across nodes, as well as predict improvements in runtime based on the predicted partitioning.

The auto query plan improvement capability relies on the optimizer engine to learn and improve the query plan based on queries executed earlier, powering the vital ability to become more intelligent over time. In contrast, traditional caching techniques are not intelligent, thereby constricting query agility for customers. For instance, Autopilot can improve TPC-H, TPC-DS 24TB performance by up to 40%, further validating the MySQL HeatWave architectural design strategy.

The auto scheduling function directly enables reduced query wait time for mixed OLTP/OLAP workloads. Since analytic queries take longer than OLTP queries, HeatWave directly addresses the issue by predicting the execution time for each query. Then, short queries, say from an online commerce application that require fast turnaround, are prioritized over long running queries, such as reports where time is not as critical, enabling the system to reduce wait time for shorter queries (and keeping customers happy) without changing the total execution time.

Scale-Out Data Management for MySQL HeatWave

With the introduction of MySQL HeatWave, customers are storing more data than even before in the MySQL database. As the size of the data grows, there is a need to manage this data more efficiently and specifically to optimize the time it takes to load data from the MySQL database into the HeatWave cluster. To meet this growing demand, Oracle is offering MySQL HeatWave Scale-Out Data Management to accelerate the data reloading time to HeatWave. With this capability any amount of data can be reloaded into HeatWave in constant time. To achieve this end, data is portioned and stored in the object store using an encrypted, in-memory format. This approach means data can be reloaded in parallel by multiple nodes of the HeatWave cluster at object store bandwidth. For a 10TB data size for instance, this results in a 100x improvement in reload time.

Now customers can reload their data as needed to restart, upgrade, or recover from errors with finer granularity, along with having changes to MySQL propagated to the object store. Of equal importance, the data is always encrypted—without any impact on performance.

Moreover, MySQL HeatWave now scales up to 64 nodes and can process up to 32TB of data, and improved scalability is achieved as the number of nodes increase by 20% over the earlier release.

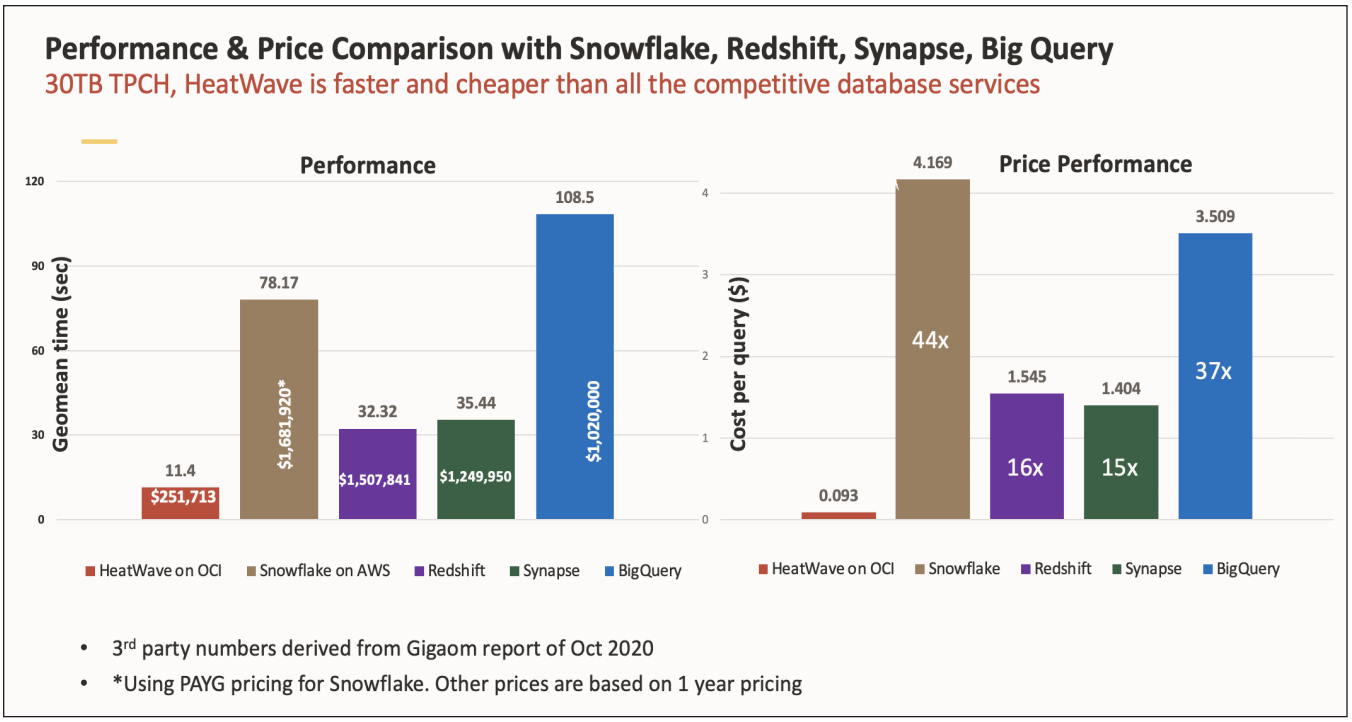
... customers can reload their data as needed to restart, upgrade, or recover from errors with finer granularity, along with having changes to MySQL propagated to the object store.

MySQL Database Service with Heatwave: Let the Price/Performance Competition Begin

The core capabilities of MySQL HeatWave allow Oracle to demonstrate the most compelling price and performance advantages against its main competitors. For example, against Amazon Aurora - based on benchmark queries that are derived from the industry standard TPC-H benchmark - the MySQL Heatwave offering, based on 4TB TPCH, delivers a 6.3 second GeoMean of query time in comparison to Aurora's 2.5 hours, 1400x faster. In addition, HeatWave's annual costs, based on 10 E3 nodes, are $34,073, which is half the price compared to Aurora's calculated $67,336 annual cost, both based on list pricing.

Likewise, MySQL HeatWave delivers comparable advantages against Amazon's RDS MySQL offering as the 6.3 second GeoMean of query time is a remarkable 5400x faster than Amazon's RDS MySQL's 11 hours GeoMean of query time and the MySQL HeatWave's $34,073 annual cost, based on 10 E3 nodes, is @2/3 the cost of the RDS MySQL annual cost of $54,393.

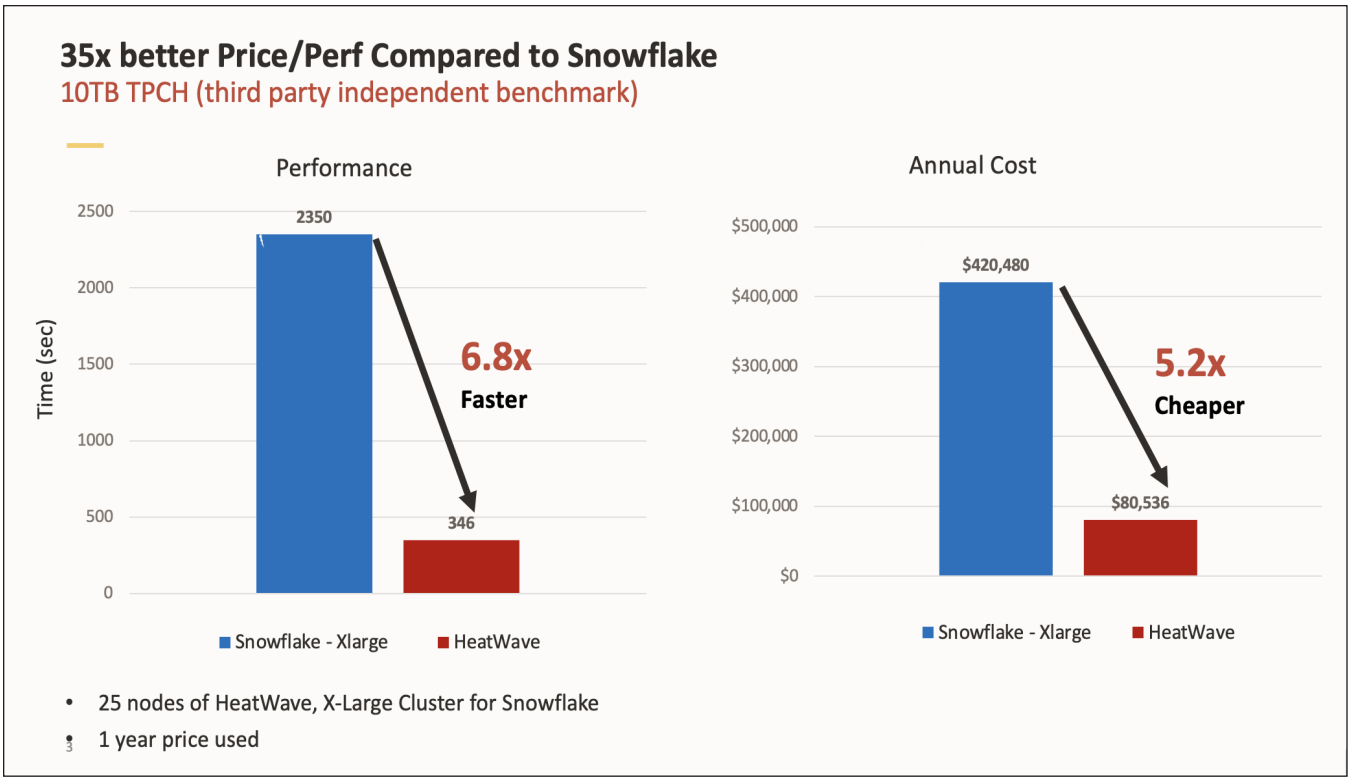
The price/performance advantage becomes even more pronounced when comparing MySQL HeatWave on OCI against a wider array of competitors including Snowflake on AWS, Redshift, Azure Synapse, and Google BigQuery as shown in this chart:



Taken together, we view the entire chain of new MySQL HeatWave capabilities producing another set of impressive price/performance advantages that further validate Oracle's competitive differentiators. For example, MySQL HeatWave yields an overall thirteen-fold price/performance edge compared to Redshift AQUA as HeatWave requires only 346.5 seconds to Redshift's 2380 seconds to complete the TPC-H benchmark,6.8x faster performance, while costing 1.9x less with an $80,536 calculated annual cost compared to Redshift's $150,784 annual cost (based on the 10TB TPC-H third party independent benchmark using 25 nodes of HeatWave, 4 nodes of RA3.4xLarge). Also of note, we see AWS position AQUA as an "advanced" query accelerator but the only thing it seems to accelerate is the rate at which customers spend on Redshift.

What would happen to the price/performance differential if a customer chose to cut annual costs for its Redshift AQUA service? We see the contrast remains stark as HeatWave continues to require only 346.5 seconds but now Redshift needs 6310 seconds, generating an 18x faster performance edge for HeatWave while costs are comparable based on Redshift's $75,392 annual cost and HeatWave's $80,536 annual cost. The trade-off in performance vs. cost is immense and one that no organization should be willing to risk.

Casting the net out further, how does MySQL HeatWave stack up against Snowflake (based on the 10TB TPC-H third party independent benchmark using 25 nodes of HeatWave, X-Large Cluster for Snowflake)? We see that the price/performance gap increases significantly to an overall 35x differential. This includes HeatWave's having a 6.8x faster performance edge, requiring only 346 seconds in relation to Snowflake's 2350 seconds, and a 5.2x cost edge shown by HeatWave's $80,536 annual cost compared to Snowflake's $420,480 annual cost. We view Snowflake costs as extremely high as the company forces users to select "T-shirt/shoe sizes," in increments of 16, 32 or 128 nodes. So, for example, if a customer needs 18 nodes they are forced to buy 32 nodes. This notable lack of granularity makes as much sense as going into Foot Locker and having to walk out with a size 22 shoe when you have size 10 feet.



In addition, MySQL HeatWave yields an overall thirteen-fold price/performance edge compared to Redshift AQUA as HeatWave requires only 346.5 seconds to Redshift's 2380 second, 6.8x faster performance, while costing 1.9x less with an $80,536 calculated annual cost compared to Redshift's $150,784 annual cost (based on the 10TB TPC-H independent benchmark).

In the interest of full transparency, Oracle is making the latest benchmark code and scripts available on GitHub and Oracle.com, enabling customers to perform them using their own data. This is key, as most benchmarks from vendors are clouded in secrecy, and Oracle is asking anyone in the industry to replicate their benchmarks as often as they like. We view this approach as a fully transparent marketing strategy, as opposed to some approaches where vendors throw up diagrams with no hard, factual data.

MySQL Database Service with HeatWave: Why Window Function Support is Key

Window functions are popular in many analytic applications. Window functions were introduced in MySQL in version 8.0. With this release, HeatWave has added support for window functions so that queries with window functions can be accelerated by HeatWave.

The implementation of window functions in HeatWave leverages the massively partitioned architecture of HeatWave. Data is partitioned across the various cores across the different nodes of the HeatWave cluster. The processing of window functions is done within a core in parallel with all the cores using a state-of-the-art algorithm. As a result, there is no need for any global sorting operation.

This approach results in the ability to perform window function computation for each tuple. Now all cores can compute in parallel using removable cumulative aggregation that results in most aggregates computing in O(N) time complexity as well as minimum and maximum in O(N logN) mode. We identify HeatWave's highly parallel and partitioned architecture as vital in powering Oracle's ability to separate its MySQL HeatWave solution from competitive offerings.

MySQL Database Service with HeatWave: Turbo-boosting both OLTP and OLAP Workloads

Through cloud-first auto scheduling innovation, MySQL HeatWave offers the topmost performance in the MySQL segment including in comparison to Amazon Aurora for both OLTP and OLAP. To review, changes made to the MySQL database are propagated to HeatWave memory and scale-out storage functions, enabling complete transparency without any need for human intervention. This means that queries are always conducted across the latest data.

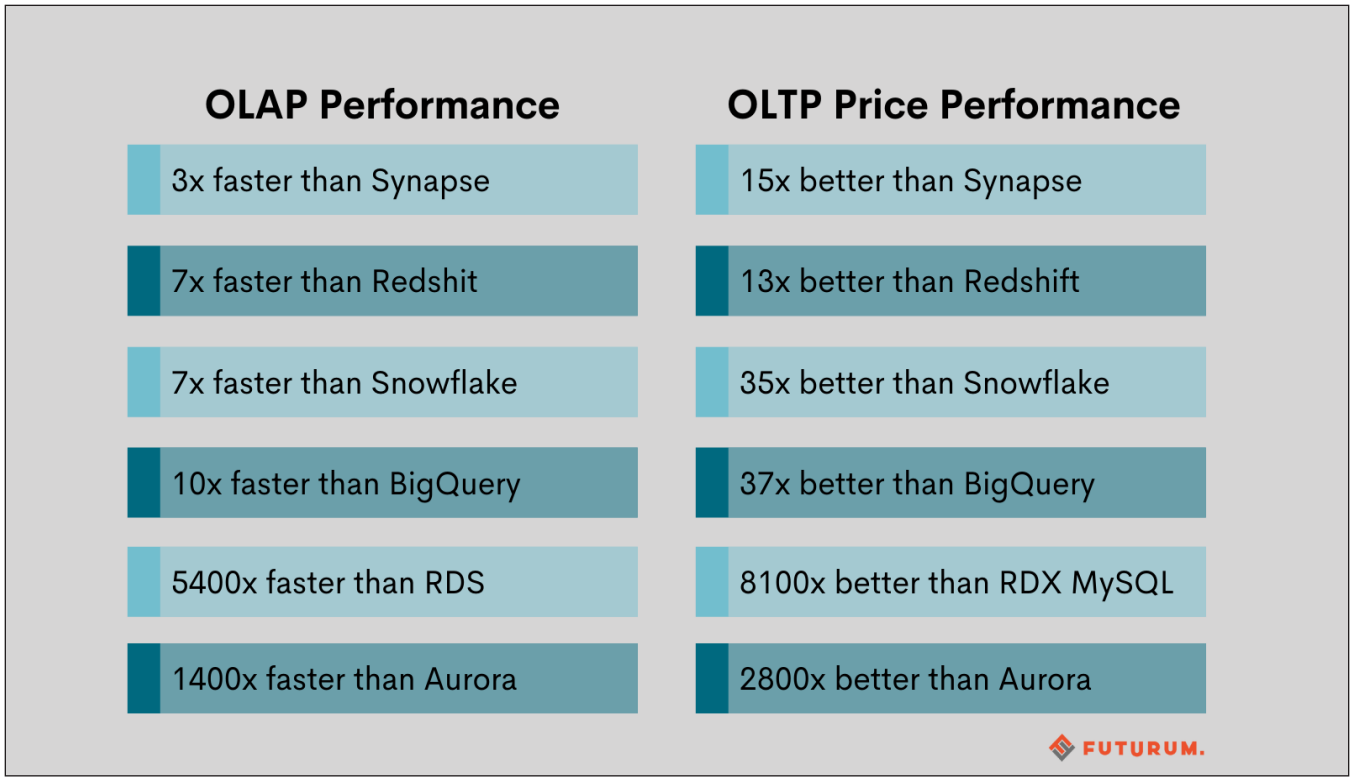
We view using the CH-benCHmark framework as providing the benchmark readings needed to fully understand mixed OLTP/OLAP workload outcomes. This mixed workload benchmark blends industry standard TPC-C (Transaction Processing Performance Council) and TPC-H (Transaction Process Council Ad-hoc/decision support) benchmarks, which are essential to simulating real-world applications which have a mix of both OLTP transactions and OLAP queries. Through the benchmark process, the TPC-C schema and statements are kept as is, while TPC-H queries and schema are modified. In conducting the benchmark, the net workload consists of the 5 TPC-C transactions and 22 queries adapted from TPC-H, allowing the benchmark to measure throughput and latency for OLTP and OLAP.

We note that for mixed workloads, MySQL Heatwave registers substantially faster than Amazon Aurora OLAP across four concurrent OLAP sessions (CH-benCHmark, 100G), consisting of 18x lower latency with Heatwave only taking 35 seconds while Aurora (db.r5.8xlarge) takes all the way up to 632 seconds; 110x higher throughput for HeatWave delivering .06 Trnx/minute but Aurora taking 6.6 Trnx/minute; and 2.4x cheaper as HeatWave's calculated annual costs are only $9,292 while Aurora is an expensive $22,162.

For OLTP only workloads, MySQL HeatWave delivers the same performance as Aurora for OLTP at half the cost (based on the CH-benCHmark, 100G, TPCC, 128 concurrent sessions, 30K OLTP transactions/min). With both HeatWave and Aurora running at the same latency of .02 seconds and throughput (Trnx/minute) of 30K. HeatWave's annual costs are only $9,292 while Aurora is more than double the cost at $22,162.

Upon combining the OLTP and OLAP differentials, MySQL HeatWave comes in at an astonishing 20x-500x faster and 50x-1200x better price/performance ratio than Aurora. With such astonishing advantages in HeatWave's favor, we see AWS' ability to close the gap in a meaningful way as an exceedingly difficult task in the foreseeable future.

MySQL HeatWave's clear OLTP and OLAP mixed workload advantages extend across the spectrum of Oracle's major rivals as the following table indicates:



Recommendations and Conclusion

MySQL Decision Makers Must Consider MySQL HeatWave. MySQL decision makers must consider the new MySQL Database Service with HeatWave solution since it delivers the best performance and price/performance metrics across latency, throughput, and annual cost considerations against the full spectrum of competitive offerings, providing Oracle virtually insurmountable advantages.

Machine Learning Assures MySQL HeatWave Automation Advantages. MySQL decision makers need to fully consider ML's integral role in ensuring that HeatWave gains more intelligence over time and to lock in new capabilities such as having a DB model that is specific to their instances and can provide explanations for recommended changes.

Evaluate the MySQL HeatWave Autopilot Architecture. MySQL decision makers must assign top consideration to MySQL Autopilot to take full advantage of ML-based automation innovations across service lifecycle, ML modeling, statistics extrapolation, and adaptive sampling capabilities.

MySQL HeatWave's support of extensive machine learning-based automation enables the entire system to become more intelligent over time, providing accrued lifecycle value for all customers. The scale-out data management capabilities assure the acceleration of many operations across any MySQL HeatWave environment. Moreover, HeatWave can now support clusters of 64 nodes and 32TB of data, while scalability when nodes are added is improved by 20%.

Of topmost consideration, MySQL HeatWave is faster than Amazon Aurora for blended OLTP and OLAP workloads, delivering an innovation breakthrough that liberates customers from traditional OLTP/OLAP workload limitations and trade-offs. Overall, we rate MySQL HeatWave as faster and less expensive than the alternative analytics services such as Snowflake and AWS Redshift AQUA, giving MySQL customers the warrant to fast track the evaluation of HeatWave's new range of innovations and validating that the dramatic performance and price/performance advantages are best suited for their evolving and expanding DB demands.

In sum, we believe that the MySQL Database Service with HeatWave offers the best value bar none across the entire MySQL DB market. From our perspective, with all the new innovations Oracle has introduced into MySQL HeatWave, the solution stands out with clear and unparalleled price/performance advantages against the gauntlet of competition. Not investing in MySQL HeatWave and instead spending twice the amount of money on other slower database cloud services such as Snowflake makes about as much sense as spending more to take a flight that gets you from London to New York in 6 days with 15 connections rather than a direct route that takes 8 hours.

The bottom line is we believe the competition just got outplayed on every measurable metric imaginable. HeatWave is the physical manifestation of nearly 10 years of deep database engineering techniques, over 5 dozen patents, and demonstrates what real cloud database innovation looks like in 2021. It also represents a wake-up call for the industry and a rude awakening to the database cloud competition as they all must now respond to the MySQL HeatWave innovation juggernaut.

Important Information About This Paper

CONTRIBUTORS:
Daniel Newman
Founding Partner + Principal Analyst, Futurum Research

Shelly Kramer
Founding Partner + Lead Analyst, Futurum Research

Ron Westfall
Research Director + Senior Analyst

PUBLISHERS:
Daniel Newman
Founding Partner + Principal Analyst, Futurum Research

Shelly Kramer
Founding Partner + Senior Analyst, Futurum Research

INQUIRIES: Contact us if you would like to discuss this report and Futurum Research will respond promptly.

CITATIONS: This paper can be cited by accredited press and analysts, but must be cited in-context, displaying author's name, author's title, and "Futurum Research." Non-press and non-analysts must receive prior written permission by Futurum Research for any citations.

DISCLOSURES: This paper was commissioned by Oracle. Futurum Research provides research, analysis, advising, and consulting to many high-tech companies, including those mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

ABOUT ORACLE
Oracle offers integrated suites of applications plus secure, autonomous infrastructure in the Oracle Cloud. For more information about Oracle (NYSE: ORCL), please visit us at **oracle.com**.

ABOUT FUTURUM RESEARCH
Futurum is an independent research, analysis, andadvisory firm, focused on digital innovation and market-disrupting technologies and trends. Every day our analysts, researchers, and advisors help business leaders from around the world anticipate tectonic shifts in their industries and leverage disruptive innovation to either gain or maintain a competitive advantage in their markets. **Read our disclaimer statement here**.

CONTACT INFORMATION

Futurum Research, LLC | futurumresearch.com | 817-480-3038 | info@futurumresearch.com
Twitter: @FuturumResearch