ORACLE UNLEASHES A HEATWAVE IN THE MYSQL MARKET

EXECUTIVE SUMMARY

In enterprise data management, Oracle Database is a dominant player, and MySQL is very popular among developers. MySQL is so prolific it ranks first in market share among all database distributions. Additionally, MySQL is the default database engine in most application stacks used in cloud-native applications.

What makes MySQL so popular is its rich functionality as an open-source relational database management system (RDMS) optimized for transaction processing. While it initially found rapid adoption in small to midsized companies, it also found a following inside the departments of large organizations.

Over many years of standardizing on MySQL, two things happened. First, many small and midsized companies that standardized on MySQL grew into large organizations that still rely on it as a data management platform. Second, departments and business units in large organizations that standardized on MySQL accumulated a large amount of critical data in this database.

Further, many departments and business units continue to adopt MySQL. This dynamic is good news for MySQL and its parent company, Oracle. It is also welcome news for the many cloud providers that have built as-a-service offerings around MySQL and the many open-source tools built around this prolific database platform.

However, for the customers that have come to rely on MySQL, there is a challenge. As an online transaction processing (OLTP) database, MySQL was not designed to enable the deep and complex analytics many organizations now require. While some companies and cloud offerings have attempted to offer some level of analytics as an add-on, none allow for real-time analytics. Until now.

While the world has been steadily increasing its usage of MySQL, Oracle has been investing in research and development to build a cloud-optimized solution to enable real-time analytics and complex OTLP queries of data residing in MySQL databases. The result is HeatWave, an in-memory query accelerator designed specifically for MySQL and integrated into Oracle MySQL Database Service. This research report will

explore HeatWave and how this offering can drive more significant insights for data-driven companies worldwide.

Disruption Is Good—Disruption Is Necessary

In any industry, over time, vendors reach a level of comfort with their existing offerings and assume a status quo position. But that stasis never lasts long. No industry is immune from disruption in the race for accelerated differentiation and material market share gains. Think of Tesla disrupting the old guard of grizzled and tired automotive vendors back in 2013 with its revolutionary electric cars, or AWS disrupting the legacy hardware vendors back in 2006 with the fundamental advent of infrastructure as a service (IaaS).Disruption is a reliable event.

The MySQL Database community is no different. As the installed base has continued to grow, there has been little to no major innovation for the past 10 years. Vendors have forked off open-source versions of MySQL and hosted it on their cloud, but that's about the same as parking your car in a skyscraper (or putting your car into orbit) and calling it an innovation—you've just moved to a different location while the underlying car technology has remained the same.

While established MySQL cloud database providers such as Azure, AWS, and Google have all tinkered with their offerings over the years, there have been no real cost and performance breakthroughs until recently, from the company one would expect: Oracle.

In a Data-Driven World, Analytics Is Key

To say data-based insights drive modern business is an understatement. Refined real-time data—meaning data transformed into intelligence through analytics engines that can deliver deep insights and better actionable decisions—drives the successful business.

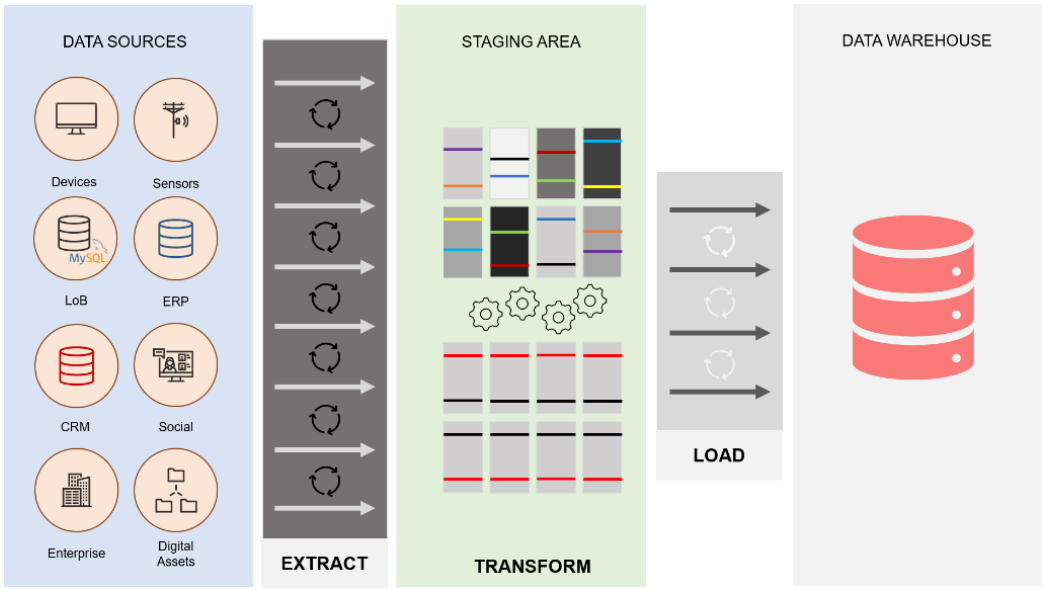
The challenge associated with driving accurate insights in real time is multi-dimensional, and the amount and type of data generated are two significant factors, among others.

In the case of complex queries on large datasets, the first issue is time—the time it takes an analytics engine to execute a query and return a valuable result to the requesting entity. The second issue is equally complex. In any data-driven organization, data resides in different sources in departments and business units, and these data sources are often MySQL databases. As any database manager with experience

knows, MySQL was not designed to execute complex queries across multiple OLTP databases. Further, no one developed a tool to perform online analytical processing (OLAP) across numerous MySQL databases.

As a result of this challenge, organizations turned to more manual, multi-step methods to perform OLAP queries for MySQL environments. Often extract, transform & load (ETL) tools such as Xplenty or Oracle Data Integrator may take data from multiple MySQL sources and house it in an OLAP destination, such as a data warehouse, and, finally, run analytics on the data.

FIGURE 1: EXTRACT, TRANSFORM & LOAD PROCESS



Source: Moor Insights & Strategy

But there are several issues with this method of managing your MySQL environment:

1. **Data staleness.** The ETL process takes time, with a cycle taking upward of four hours. This can lead to reporting and insights that are based on stale data, which can significantly impact all facets of business operations.
2. **Security.** As the value of data has increased significantly, hackers are constantly looking for ways to capture and hold data for ransom. While ETL tools make security a priority, any time data is in flight information is at risk as it presents another attack surface for hackers.
3. **Maintaining multiple database environments and analytics tools.** Database environments and tools are expensive, directly and indirectly—from licensing

costs to infrastructure to the experts required to manage each semantic and overall environment.

In short, real-time analytics for MySQL has been anything but real-time because the data in an OLAP environment is out of sync with the OLTP environment.

WHAT ABOUT THE CLOUD?

Do services such as Aurora, Redshift, Snowflake, or Synapse enable organizations to achieve the same goals? Yes and no. In theory, such services can assist organizations in running analytics against recent data. But due to the architecture of these solutions, such analysis always requires ETL to move data to separate databases and tools (some cloud providers have several for customers to choose from, further complicating the process). This is both time-consuming and costly and fails to deliver real-time analytics.

The challenge of running real-time analytics on MySQL environments has a tangible impact on the productivity of companies of all sizes. Moor Insights & Strategy sees resolving this issue as a potential game-changer for the continued enterprise growth of MySQL by removing a critical barrier to adoption.

HEATWAVE—THE GAME-CHANGER

For organizations seeking real-time analysis of their MySQL environments, it makes sense that they would look to the most invested company in MySQL. Oracle announced the availability of HeatWave, a high-performance in-memory query accelerator for MySQL Database Service, in late 2020. HeatWave is an analytics service available through Oracle Cloud Infrastructure (OCI) that can deliver real-time analysis on real-time data faster and cheaper than any other service, according to Oracle. Based on the performance data that Moor Insights & Strategy has reviewed, this appears to be an understated way of describing HeatWave's capabilities. Moor Insights & Strategy has not seen another unified solution that addresses transaction processing and analytics in the MySQL market to date—meaning that HeatWave's innovation is clearly in a category of one.

Unlike other MySQL cloud database services on the market today, Oracle MySQL Database Service with HeatWave can run transaction processing and analytics without the need for any data movement, a first for this space. Customers can also make their decisions in real-time with HeatWave because there is no ETLing the data—all the data is inside a single database. As soon as organizations make any changes, they are

visible for queries immediately, which is quite different from current MySQL database cloud services.

IT'S ALL IN THE ARCHITECTURE

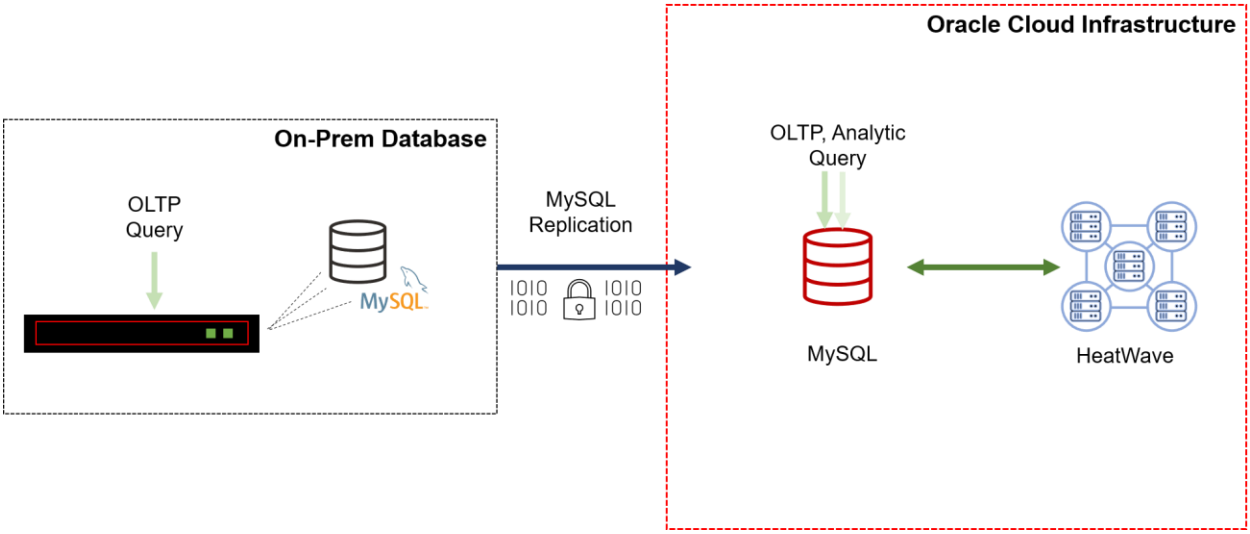
The two keys to successfully executing real-time analytics of a MySQL environment (or environments) are:

- Ensuring that the data is continuously fresh
- Executing the query as fast as possible

While vendors have been delivering analytics tools to MySQL users in a somewhat klugey fashion, Oracle has been deep at work in Oracle Labs. The company quietly developed HeatWave – a cloud-native massively scalable service that can handle real-time analytics of an organization's MySQL data regardless of where that data resides – in the cloud or on-premises.

Using the MySQL Database Service and HeatWave, data inserted into MySQL moves into HeatWave in seconds. If an organization runs MySQL on-premises, those databases perform a real-time replication with the HeatWave service in OCI.

FIGURE 2: ON-PREMISES ANALYTICS WITH HEATWAVE

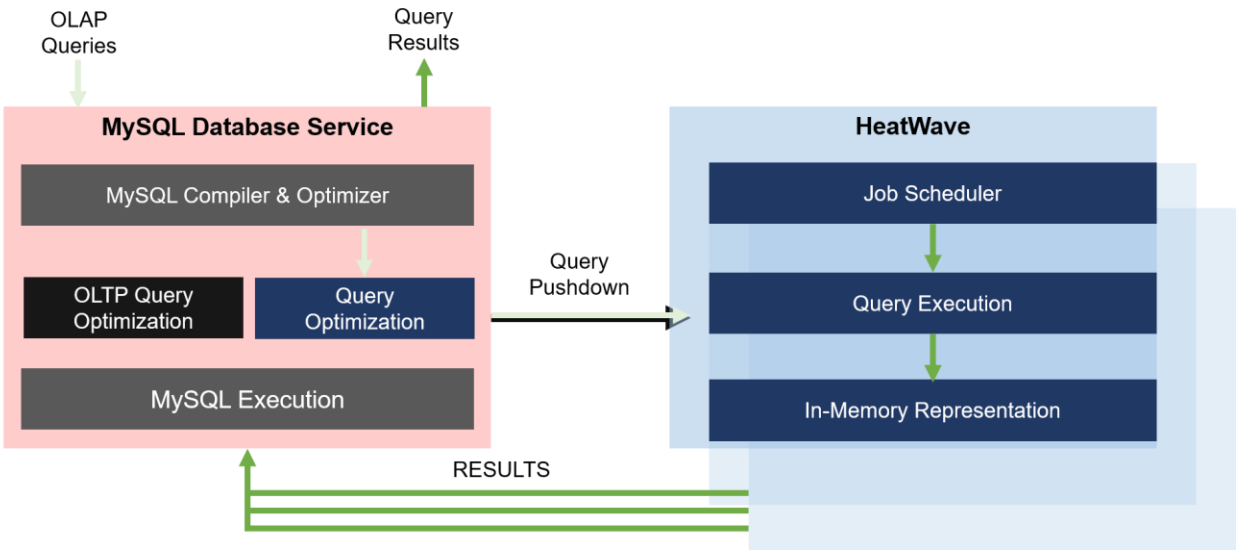


Source: Moor Insights & Strategy

As queries are sent to the MySQL database for execution, the MySQL optimizer makes a cost-based determination about the appropriate engine for execution. Simple OLTP

queries are handled as a normal process and passed to the MySQL engine. In contrast, analytics or complex queries are sent to HeatWave, which has an optimized in-memory representation of the data, and the query execution is performed in a highly parallelized fashion.

FIGURE 3: REAL-TIME ANALYTICS PROCESS IN HEATWAVE



Source: Moor Insights & Strategy

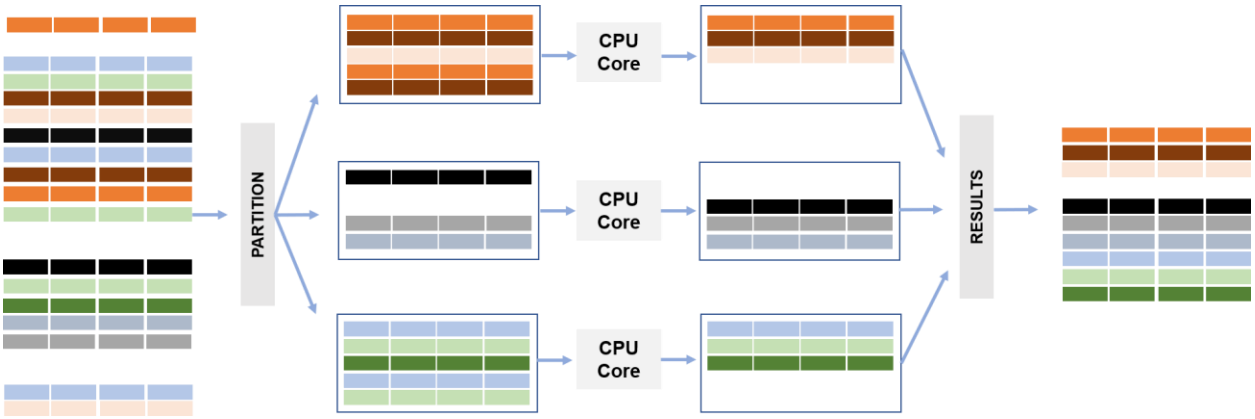
HEATWAVE—DESIGNED FOR THE CLOUD FROM DAY ONE

It's worth spending a few words describing how Oracle designed HeatWave for the cloud with an eye toward performance, scale, and cost—areas it highlights as key differentiators. The company achieved this through three design decisions:

1. ***In-memory, hybrid columnar processing***. The data is loaded in the memory of the HeatWave engine in a hybrid columnar format optimized for vector execution.
2. ***Massive inter-node and intra-node parallelism***. As queries are received, data is partitioned and spread across multiple CPU cores for query processing. This parallelized process significantly improves performance—enabling HeatWave to return results drastically faster than other services.
3. ***New algorithms for distributed query processing.*** The Oracle Labs team has invented new algorithms for many of the key analytics operators such that they scale to a large number of servers and can be optimized using the hardware instructions of the underlying machines in the OCI. These algorithms help

HeatWave take advantage of commoditized hardware and object storage, contributing to both higher performance and lower cost.

FIGURE 4: MASSIVE PARALLELISM



Source: Moor Insights & Strategy

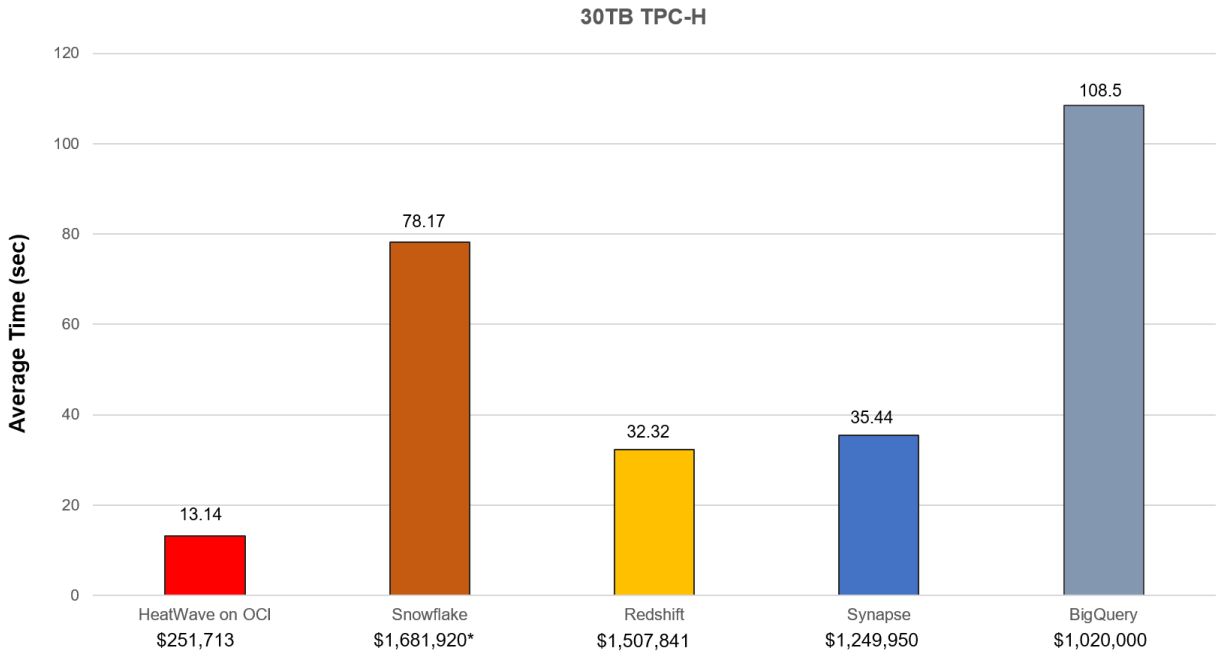
The process by which HeatWave ingests, parses, and analyzes OLTP data can seem both impressive and complex, and it is. However, what makes HeatWave most attractive to Moor Insights & Strategy is how easily it is consumed. Using HeatWave is as simple as flipping the proverbial switch—no modifications or re-factoring of applications, no database re-architecting, no user intervention or setting changes required. Oracle has abstracted all of the complexity from users and applications—true consumerization of data analytics.

WHAT MAKES HEATWAVE UNIQUE—A CLOSER LOOK AT PERFORMANCE

Oracle has made very bold claims about the performance of HeatWave relative to other cloud services. And those claims are backed by public data posted to GitHub so that anyone can replicate the benchmarks using the same parameters and scripts. As previously noted, preparing and executing queries on HeatWave relative to other cloud services is not just faster and cheaper—it is so by orders of magnitude.

While Moor Insights & Strategy is usually hesitant to show vendor-driven benchmark information, these performance numbers are based on industry-standard benchmarks, have been reviewed carefully, and are legitimate. In Figures Five and Six, you can see the impact of ETL on the performance of several popular cloud-based analytics services.

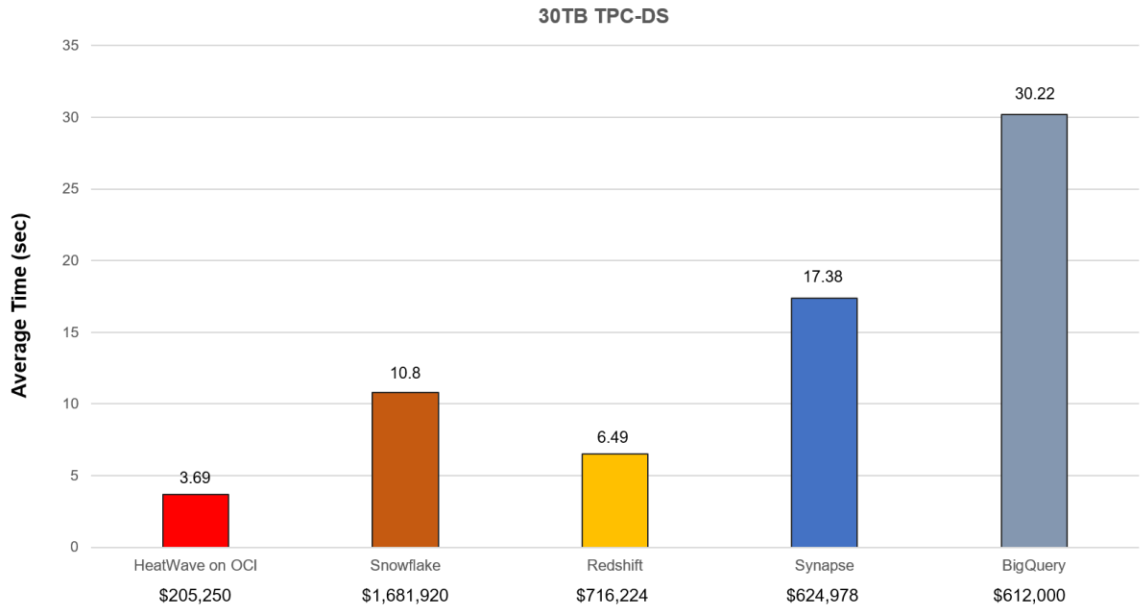
FIGURE 5: HEATWAVE TPC-H PERFORMANCE



* Snowflake pricing is based on a pay as you go model

Source: Moor Insights & Strategy/Oracle/GigaOM

FIGURE 6: HEATWAVE TPC-DS PERFORMANCE



Source: Moor Insights & Strategy/Oracle/GigaOM

Not shown in the figures above, but according to publicly available benchmarks and scripts—which, again, are on GitHub for anyone to replicate—for smaller data sizes (½ TB), MySQL HeatWave is 400 times faster than other offerings in terms of latency. Moving to a larger data size (4TB), MySQL HeatWave is 1,100 times faster than AWS Aurora and 3 times faster than the fastest shape of AWS Redshift—and at a third the cost. And Oracle ran these benchmarks on the latest, most optimized releases of AWS Redshift and Aurora, based on their guidelines and scripts.

As mentioned earlier, Oracle achieves such performance gains and cost savings through its design of HeatWave and the underlying cloud infrastructure. Object storage combined with AMD EPYC processors and an Ethernet backbone enables the OCI team to throw significant computational resources at a problem with massive parallelization.

In Closing

Managing the massive amounts of data generated every day is one of the most significant challenges IT organizations face. Moor Insights & Strategy knows this through daily conversations with CIOs and IT executives.

An equally vexing challenge is extracting the volumes of data residing in the departmental and business unit databases running MySQL. Databases that have been growing year after year with critical data are necessary to make informed business decisions.

The response of many solutions providers to date has been bolt-on approaches that utilize some method of ETL to move data to an analytics tool or service while trying to analyze the freshest set of data possible. This can lead to stale data that can, in turn, misinform the business.

Virtually every major cloud provider tries to achieve real-time analytics through a series of specialized services requiring data to be stored in its cloud and utilizing time-consuming and expensive methods to move and analyze data.

As the custodian of MySQL, Oracle has delivered a real-time analytics engine designed from the ground up to take advantage of cloud architectures. HeatWave is the only natively integrated analytics engine for MySQL and is unmatched in the industry in terms of performance and cost.

In addition to HeatWave's performance, it is a service that is very consumable—both in ease of use and cost. MySQL databases on-premises can use HeatWave by replicating between their on-premises databases and the MySQL Database Service with HeatWave in OCI. No modifications to the MySQL environment or supported applications are required. Further, OCI's use of commodity hardware reduces cost and enables a scaling that accelerates the parallelized function of HeatWave's analytics.

Oracle stands alone in offering a unified solution for both transaction processing and real-time analytics for MySQL. While others may claim such capabilities, requirements for ETL make such services anything but real-time in nature. And Oracle's strategy of pricing HeatWave instances cheaper than OCI compute instances signals an aggressive approach to customer acquisition. As a result, organizations that deploy MySQL and consider the data stored in their MySQL environment critical to business should evaluate Oracle MySQL Database Service with HeatWave as a top priority.

To learn more about Oracle's HeatWave service, visit https://www.oracle.com/mysql/heatwave/

IMPORTANT INFORMATION ABOUT THIS PAPER

CONTRIBUTOR

Matt Kimball, Senior Analyst at Moor Insights & Strategy

PUBLISHER

Patrick Moorhead, Founder, President, & Principal Analyst at Moor Insights & Strategy

INQUIRIES

Contact us if you would like to discuss this report, and Moor Insights & Strategy will respond promptly.

CITATIONS

This paper can be cited by accredited press and analysts but must be mentioned in context, displaying the author's name, author's title, and "Moor Insights & Strategy." Non-press and non-analysts must receive prior written permission by Moor Insights & Strategy for any citations.

LICENSING

DISCLOSURES

Oracle commissioned this paper. Moor Insights & Strategy provides research, analysis, advising, and consulting to many high-tech companies mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

DISCLAIMER

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. Moor Insights & Strategy disclaims all warranties regarding the accuracy, completeness, or adequacy of such information and shall have no liability for errors, omissions, or inadequacies in such information. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators and not precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could cause actual results to differ materially. You are cautioned not to place undue reliance on these forecasts and forward-looking statements, which reflect our opinions only as of the date of publication for this document. Please keep in mind that we are not obligating ourselves to revise or publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.