Cloud Database Wars: Oracle MySQL HeatWave Crushes Amazon, Microsoft, Google, and Snowflake

By Bob Evans, March, 2022

As data becomes the lifeblood of the digital economy, and as modern cloud-native databases become the beating hearts of digital businesses, Oracle has jammed its MySQL HeatWave database with a slew of enhancements allowing it to blow away competitive services from AWS, Microsoft, Snowflake, and Google.
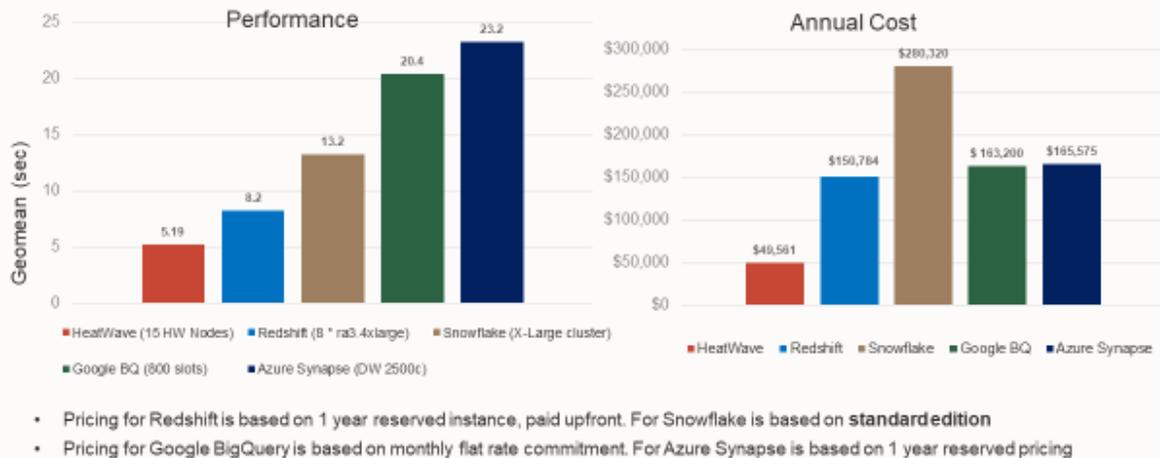That is not an opinion—it is grounded in rigorous fact-based price-performance TPCDS benchmarks, which have become a highly respected third-party standard.
Let me jump to the punchline first: in head-to-head tests showing the relative price-performance of comparable database solutions from Oracle and the four competitors mentioned above, here's how the various products stack up:

- Oracle MySQL HeatWave versus AWS Redshift: Oracle is 4.8X better
- Oracle MySQL HeatWave versus Snowflake: Oracle is 14.4X better
- Oracle MySQL HeatWave versus Microsoft Synapse: Oracle is 14.9X better
- Oracle MySQL HeatWave versus Google Big Query: Oracle is 12.9X better

(Microphone drops, clangs on floor.)

Price-Performance comparison with Redshift, Snowflake, Google BigQuery
10TB TPCDS: 4.8x better than Redshift, 14.4x better than Snowflake, 12.9x better than BigQuery, 14.9x better than Synapse

- Pricing for Redshift is based on 1 year reserved instance, paid upfront. For Snowflake is based on **standard edition**
- Pricing for Google BigQuery is based on monthly flat rate commitment. For Azure Synapse is based on 1 year reserved pricing

It can be easy to get lost in the arcana of data-management technologies and specifications and small print, particularly when they evaluate under-the-cover components that might light up the eyes of hard-core techies but don't really affect or advance the cause of the business.

Mission-critical requirements for optimizing data

That is absolutely *not* the case here, as increasingly powerful cloud databases are being deployed by more businesses for more uses and delivering more strategic value to the business than ever before. This isn't a one-time blip or spike indicative of a zany, short-lived trend; rather, we're at the beginning of the beginning, as businesses and other organizations around the globe begin to accelerate their moves to the cloud and are making cloud databases the foundations of those business and technology transformations.

Recognizing those trends, Oracle has, with this new release of MySQL HeatWave, made a huge bet that it will be able to drive massive new levels of value for customers by enhancing the traditional strengths of MySQL in transaction processing with a new set of world-class analytical and machine learning capabilities.

In that way, customers will no longer have to move their data out of a transactional system and into analytical or machine learning systems—instead, with MySQL HeatWave, a single database can handle all three functions. That fusion will save customers significant amounts of time and money, and allow them to analyze more

data and develop inferences more quickly to gain new insights that are essential to business success in today's fast-paced and fast-changing digital economy.

At the same time, this urgent demand from customers aligns with a broader trend in the market. The tech industry is responding aggressively to this heightened appetite from customers by delivering significant and ongoing advances in cloud technologies across the entire spectrum: faster and more-secure infrastructure; powerful new databases built for the demands of the digital economy; AI-powered cloud applications that give businesses better insights into their data; and industry-cloud applications that help companies master the new capabilities and processes required in the world of digital business.

The customer value of TP + analytics + machine learning in a single database

Today, every company in every industry in every part of the world is racing to:

- develop data-driven business processes with the speed and intelligence to meet the soaring needs of today's accelerated economy;
- develop data-driven businesses that enhance and extend physical or digital products that customers already know, trust, and invest in;
- develop data-powered inferences and insights that allow leaders to make better, smarter, faster, and revenue-enhancing decisions; and
- develop end-to-end data cultures that fuel the final stages of digital transformation to truly digital businesses.

None of that is possible without world-class cloud databases. And without question, Amazon and its AWS Redshift, Microsoft Synapse, Google's Big Query, and Snowflake's high-flying data cloud have all experienced impressive growth to date as some of the premier solutions in the market.

But no company on Earth knows more about enterprise-scale databases and the complexities of multifaceted data management than Oracle, and its just-completed multiple enhancements to its open-source MySQL HeatWave—a product that was pretty darned impressive in its own right before this latest round of big-time improvements—is going to change the competitive environment radically and quickly for AWS, Microsoft, Snowflake, and Google.

And of all the new features and power within MySQL HeatWave, the biggest one is that customers can now use MySQL HeatWave to handle the trifecta of transaction processing, analytical query processing and machine learning in one single and thoroughly modern cloud database. That clearly represents a triple threat to Snowflake, AWS, Microsoft and Google.

Oracle's 8-point master plan for MySQL HeatWave

Here's Oracle's master plan for how its newly enhanced MySQL HeatWave database will disrupt and reframe the competitive dynamics for cloud databases:

1. offering a single database service that handles transactions, analytics, and machine learning (Amazon Aurora is a transactional database, and must be paired with Amazon Redshift for analytics or Snowflake on AWS); Amazon Redshift ML is paired with SageMaker for machine learning; Snowflake users need a third-party tool or service for machine learning since it has no in-database ML;
2. optimizing the impact of machine learning across multiple facets of MySQL HeatWave, such as Autopilot to automate management and accelerate queries or HeatWave ML to automate and speed machine learning applications;
3. delivering security capabilities that competitors with less experience in both databases and security cannot match;
4. providing customers with ML solutions that are "explainable," an attribute that's become incredibly essential in today's economy;
5. making the database and its ML features self-tuning, which enables continuous improvement and frees up personnel to focus on higher-value work;
6. engineering MySQL HeatWave to be both faster and cheaper than competitive solutions, as noted above;
7. ensuring that the database is fully available as it scales up or, which is increasingly important as volumes of data continue to soar; and
8. software innovations which further lower the cost of running a workload with MySQL HeatWave by up to 50% and providing predictable costs in advance for customers so they can be fully confident in their financial commitments.

MySQL HeatWave has Larry Ellison's full support and backing
Before taking a more in-depth look at what Oracle did in each of those 8 areas, I also want to mention a related issue that, while not technical in nature, is nevertheless an extremely important factor in the market-share battle Oracle is launching with the new and greatly improved MySQL HeatWave database. And that related factor is the intense commitment that Oracle Founder, Chairman and CTO Larry Ellison has made to MySQL HeatWave. Oracle has a staggeringly large portfolio of cloud products, and while Ellison no doubt loves them all, he surely loves some more than others.
And judging by some recent and very high-profile comments he's made about MySQL HeatWave, it's perfectly clear that the guy who's not only CTO of Oracle but also owns 26% of the company is unconditionally committed to making it a huge success in the marketplace.

The significance of that commitment can't be overstated because, in true Ellison fashion, he has chosen to go directly at the biggest and most-successful players in the market. And in this case, those competitors include 3 of the world's largest, wealthiest, and most-powerful companies—Amazon, Microsoft, and Google—as well as high-flying relative newcomer Snowflake, which last quarter saw its revenue increase by 106%, albeit with an operating loss of $152.0 million.

Oracle's unparalleled expertise and experience in database technology
But for all the assets those fine companies have, Oracle's got one thing they haven't got: 40+ years of unmatched experience, expertise, research, patents, and insights into database technology and how it is used within midsized businesses and organizations on up through the largest and most-demanding corporations in the world.
And in fusing breakthrough analytical and machine learning capabilities into the existing transaction-processing strengths of MySQL, Oracle has brought to bear in MySQL HeatWave all of the company's deep technical knowledge, its unmatched levels and breadth of advanced R&D, and its experience in working with hundreds of thousands of customers deploying every sort of workload imaginable in every conceivable type of IT environment.
So during his opening remarks on Oracle's March 10 earnings call, when Larry Ellison brought up MySQL HeatWave and said, "We've never had a product so well received by customers and analysts in our history," that wasn't marketing fluff. Rather, it was Ellison taking one of his trademark competitive positions that fly directly in the face of current market reality and conventional wisdom and put the spotlight on Ellison and his company to succeed and earn the fruits of that victory, or to fail very publicly and very visibly.
And during his four decades of leading Oracle, Ellison has come out on the winning side of those all-in bets much more often than not:
- 20 years ago, when Oracle jumped into the applications, "conventional wisdom" said Oracle's a database company and doesn't know anything about applications;
- 13 years ago when Ellison acquired Sun Microsystems and the so-called experts howled that Oracle doesn't know anything about the hardware business;
- 5 years ago when Ellison announced the Autonomous Database, and the wizards of smart said such a product is too complex and will never work;
- 3-4 years ago when Ellison began to say that Oracle's Fusion ERP would become the world's leading cloud ERP solution; and

- 3 years ago when Ellison pushed Oracle into the public-cloud infrastructure business, and the geniuses all said it was a fool's errand to try to compete with the trillion-dollar trio of Amazon, Microsoft, and Google.

Ellison: 'We're going after the AWS Aurora plus Redshift business'
And, bringing this all back to MySQL HeatWave, I suspect that more than a few of those same self-described experts howled in indignation when Ellison, on that March 10 earnings call, said this about his plans to take the fight directly to market-leader Amazon:
"For Amazon, the combination of AWS MySQL Aurora and Redshift is a multibillion-dollar business. *And now we're going after that business in two ways* [emphasis added]," Ellison said.
"MySQL HeatWave is good at both transaction processing and query processing. So MySQL HeatWave doesn't simply replace Aurora—it replaces both Aurora and Redshift, or it replaces both Aurora and Snowflake, because MySQL HeatWave does transaction processing very well, replacing Aurora. And it does query processing a lot better than Redshift or Snowflake."
Okay—it's certainly fair to say those are fighting words. So the question before the house is this: can Oracle and MySQL HeatWave back up Ellison's provocative talk? Can it outperform the formidable AWS Aurora-Redshift combo so completely and thoroughly that customers will be willing and perhaps eager to go through the nontrivial task of database migration to reap the 4.8X better price-performance that MySQL HeatWave offers? Will business leaders look at that 4.8X differential and mandate that the change be made to give them the greatest chance of optimizing their data-powered digital futures?
Migration is almost instantaneous
I think there's a very good chance they'll be willing to give that 4.8X delta a lot of consideration because, setting aside the benchmark-related technological superiority of MySQL HeatWave, Oracle has also fully acknowledged customer preferences by making it a multicloud database. That gives customers the ability to run MySQL HeatWave not only in the Oracle Cloud but also within AWS and Azure.
Oracle did that, Ellison said, because it is fully confident that once businesspeople using competitive services from Amazon or Microsoft are able to see what MySQL HeatWave can do, they'll be eager to make the switch.
Ellison's rationale for the multicloud capability: "We did that because we're going after the Aurora user base and the Redshift user base and the Snowflake user base. We want to make it really easy to convert from Aurora and Redshift or Aurora and Snowflake to Oracle HeatWave. And if we're running on AWS, for example, you

press a couple of buttons, and your data is moved immediately to Oracle MySQL HeatWave."

Emphasizing that "press a couple of buttons" was not a catchphrase for a process that's actually much more complex and time-consuming, Ellison said that Oracle has built into MySQL HeatWave almost all of the code required to make the migration mostly invisible to customers so that they can immediately begin deriving value out of the new database.

"You do not have to change your application at all," Ellison said, "and the cost-performance benefits of moving to MySQL HeatWave are extraordinary."

Customer feedback driving enhancements to MySQL HeatWave
The MySQL HeatWave development team is led by Oracle senior vice-president Nipun Agarwal, who said that while Oracle had many of its own ideas about how to optimize the performance, flexibility, and customer focus of the new database, the team also incorporated large volumes of customer input into its plans.

That direct input came from customers migrating to MySQL HeatWave from a variety of databases, including three from AWS—Aurora, Redshift and RDS—as well as Teradata, SAP HANA, and Google BigQuery.

Agarwal outlined four particular enhancements of greatest interest to those customers:

1. More machine-learning capabilities. As the volumes of data being stored in MySQL HeatWave rise, customers wanted Oracle to boost the ML capabilities within the database.

2. Real-time elasticity. While MySQL HeatWave is highly scalable, customers asked Oracle to automate and eliminate downtime during resize of the HeatWave cluster.

3. New entry-level pricing. Some customers wanted to trade-off some performance for an even lower entry-level price point.

4. New benchmarks. Several customers wanted Oracle to run different types of benchmarks so those customers could better understand competitive capabilities across different types of workloads.

8-point plan: how Oracle made MySQL HeatWave 4.8X better in price-performance than Amazon Redshift
Building on the 8-point master plan for MySQL HeatWave noted above, let's take a look in more detail at each of the advantages called out by Oracle:

1. offering a single database service that handles both transactions and analytics (AWS Aurora is a transactional database, and must be paired with AWS Redshift or Snowflake on AWS for analytics);
2. optimizing the impact of Machine Learning across multiple facets of MySQL HeatWave;
3. delivering security capabilities that competitors with less experience in both databases and security cannot match;
4. providing customers with ML solutions that are "explainable," an attribute that's become incredibly essential in today's increasingly ML-driven economy;
5. making the database and its ML features self-tuning, which enables continuous improvement and frees up personnel to focus on higher-value work;
6. engineering MySQL HeatWave to be both faster and cheaper than competitive solutions, as noted above;
7. ensuring that the database is fully available as it scales the cluster, which is increasingly important as volumes of data continue to soar; and
8. software innovations which further lower the cost of running a workload with MySQL HeatWave by up to 50% and providing predictable costs in advance for customers so that they can be fully confident in their commitments.

1. Single database service for transactions, analytics and machine learning
The benefits of Schumpeter's "creative destruction" are seen every day in the tech business, and the rise of the cloud has only intensified that trend. The old model that worked so well for six decades since IBM System 360—an entire database stack over here for transactions, a separate one over there for analytics, and a separate one for machine learning—seems wildly out of place in the context of today's data-driven business world. The resulting need to move data out of one system and into the other to gain the full value of that data has become too expensive, too time-consuming, too labor-intensive, and too fraught with security exposure to endure. One MySQL HeatWave database is simpler than two from AWS plus all the associated ETL tools and data movers—that's a fact, not an opinion.
For the past decade, we've heard about new types of databases that are said to have the capability to do both transaction processing and analytics together. But so far that is still the exception rather than the rule. And this is why Larry Ellison is so bullish on MySQL HeatWave: it pulls both essential functions into a single system so that businesses spend less time and money delivering greater value to their customers and, consequently, to their own top and bottom lines as well.

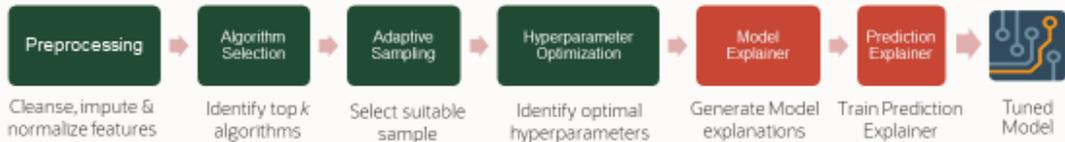**2. Optimizing ML processing with HeatWave Machine Learning**
Other MySQL versions require the customer to perform the complex chores involved in ETL: extracting data from the transactional system, transforming into formats that work with the analytical or machine learning system, and then loading the data into the appropriate system. These antiquated processes are time-consuming, labor-intensive, expensive, and—as if all that's not bad enough—security risks. In a world that is becoming heavily dependent on data, data analysis, and data insights, that's a recipe for disaster because all that time and complexity prevents businesses from being able to make decisions that move at the speed of their customers.

But with HeatWave, ETL goes the way of the dodo bird and the data can be analyzed and machine learning models can be generated in the same database in which the transactions were posted in real-time. For customers, that means eliminating security exposure, less labor, and lower expenses, and more speed. And all of that rolls up to mean significantly more value for businesses as it gives them the ability to have fresh insights about their customers and move as fast as they are every step of the way.

The new HeatWave ML automates the 3 critical steps of training, inference, and explanation without any intervention from the user, which gives customers the full benefit of powerful new ML technologies without having to develop advanced skills in this highly advanced and fast-changing field. And this is one of the areas in which Oracle's decades of expertise in related areas comes into play: HeatWave ML accelerates the training phase from hours to seconds, Agarwal said.
 (See graphic below.)

Training with HeatWave ML is fully automated
Several patents for improving performance and accuracy

Preprocessing → Algorithm Selection → Adaptive Sampling → Hyperparameter Optimization → Model Explainer → Prediction Explainer → Tuned Model

Cleanse, impute & normalize features | Identify top *k* algorithms | Select suitable sample | Identify optimal hyperparameters | Generate Model explanations | Train Prediction Explainer | Tuned Model

- *Meta-learned proxy models* make accurate one-pass decision at every pipeline stage
  - Iteration-free pipeline
  - Early algorithm selection to enable accurate sampling and feature selection
- Skew aware *adaptive sampling*
- Highly parallel *gradient-based search space reduction* for hyperparameter tuning
- *Automatically converge* without relying on user to specify time budget
- All models created by HeatWave ML can be *explained*

3. Enhanced security for dangerous times.
As noted above, the elimination of the ETL process dramatically limits the exposure that companies face when moving data frequently between systems. In today's climate of increasingly aggressive and sophisticated cybercriminals, leading companies are finding new and better ways to totally bypass such exposure and, as a result, boost their cybersecurity capabilities. Over the past decade or so, the main criticisms of ETL have been the costs involved and the manual labor required to support the required processes. But now that cybersecurity has become a board-level issue, MySQL HeatWave's ability to lower security exposure by eliminating ETL has become a major attraction for customers.

4. Making ML models "explainable"
While machine learning has driven significant new innovations in the world of data on a variety of fronts, it has also presented some new challenges around the emerging area of "explainability." Leaders today feel that in order to keep the trust of their customers, employees and partners as the use of ML rises, those organizations must be able to clearly and distinctly explain how their ML solutions arrived at certain outcomes, how much bias was or was not involved. In that context, Oracle's HeatWave ML has been designed to generate understandable explanations to help customers with:

Regulatory compliance: may imply 'right to an explanation' for algorithms affecting users
- **Fairness:** by allowing validation that predictions are unbiased

- **Repeatability:** ensures that small changes in input do not lead to large changes in the explanation
- **Causality:** allows verification that only causal correlation between features and predictions are selected
- **Trust:** interpretable explanations encourage machine learning-based prediction

In developing HeatWave ML, Agarwal said that he and his team focused heavily on 3 areas: interpretability of outcomes, and whether humans can act on them; the quality of the explanations to ensure clarity and precision; and performance and scalability to be able to deliver explanations even with huge data sets.

5. Self-tuning offers multiple advantages

Another example of how Oracle's unmatched R&D depth and experience with all facets of data management has been its ability to make HeatWave Machine Learning completely self-tuning (Oracle has several dozen patents for HeatWave and approx. 24 patents for the HeatWave ML technology alone). By contrast, Redshift requires manual intervention to tune and update and revise its ML models—and, not surprisingly, that results in more time and budget devoted to maintenance and upkeep, and fewer opportunities for businesses to get fully aligned with what their customers want and need.

One example of this, Agarwal said, is how Redshift ML requires users to specify the desired runtime and the maximum number of cells that they intend to use—and larger values almost always lead to higher costs—but not necessarily better results. "With Redshift ML, users don't always get a better model by spending more," Agarwal said.

6. Customers spend less and get far better performance.

While there are many examples of MySQL HeatWave's superiority over Redshift—particularly with the Machine Learning tools that are driving so much innovation—one of the most striking involves the time required to train ML models. As you'll see in the two graphics below built on benchmark results for ML training, HeatWave ML is not only 25X faster and more accurate than Redshift ML, it is also at 1 percent of the cost—yes, you read that correctly, 1 percent of the cost!—of Redshift ML. That's *not* 25% faster and 97% less expensive—it's 25X and 1 percent of the cost. Check out the details in these 2 graphics:

HeatWave ML is *faster* and more accurate than Redshift ML
Default training time of Redshift ML

Dataset	Accuracy		Training Time (minutes)		Speedup
	Redshift ML	HeatWave ML	Redshift ML	HeatWave ML (2 nodes)	
Airlines	0.5	0.6524	90.00	2.71	33.21
Bank	0.8378	0.7115	90.00	3.72	24.19
CNAE-9	FAILED	0.9167	FAILED	5.91	N/A
Connect-4	0.6752	0.6970	90.00	7.13	12.62
Fashion MNIST	FAILED	0.9073	FAILED	181.85	N/A
Nomao	0.9512	0.9602	90.00	3.30	27.27
Numerai	0.5	0.5184	90.00	0.34	264.71
Higgs	0.5	0.758	90.00	68.58	1.31
Census	0.7985	0.7946	90.00	1.22	73.77
Titanic	0.9571	0.7660	90.00	0.47	191.49
CC Fraud	0.9154	0.9256	90.00	29.06	3.10
KDD Cup	FAILED	0.5	FAILED	3.55	N/A
GEOMEAN	0.712	0.754	90.00	3.561	25.271

O

HeatWave ML is *cheaper* than Redshift ML
Default training time of Redshift ML

Dataset	Cost		Cheaper
	Redshift ML Cost with 1 year plan	HeatWave ML Cost $ (2 nodes)	
Airlines	6.23	0.0479	130.03
Bank	5.68	0.0658	86.30
CNAE-9	FAILED	0.1045	N/A
Connect-4	6.18	0.1261	49.05
Fashion MNIST	FAILED	3.2151	N/A
Nomao	5.96	0.0583	102.14
Numerai	5.49	0.0060	913.49
Higgs	7.27	1.2125	5.99
Census	6.12	0.0216	283.95
Titanic	5.60	0.0083	674.32
CC Fraud	6.70	0.5138	13.03
KDD Cup	FAILED	0.0628	N/A
GEOMEAN	6.115	0.063	97.129

O

7. Scales with size of cluster

Because HeatWave ML provides *automated* tuning and training of models, it can easily scale so that its performance increases with the size of the cluster. So the larger the model is, the more rapidly HeatWave ML executes. That is due in part, Agarwal said, to the fact that each stage of the model's pipeline has *unique*

consideration for parallelism, enabling appropriate tradeoffs across accuracy, scalability and runtime.

8. Lower and Predictable costs

All of these improvements—and many others not mentioned here—have enabled Oracle to be able to remove all manual labor, compromises, and even educated guesswork that have been hallmarks of most databases and their administration. By software innovations like support for blocked bloom filters and advanced compression techniques, MySQL HeatWave has doubled the amount of data which can be processed on a HeatWave node without any impact to its compelling price performance or without impacting load times, as detailed in the graphic below.

Can run workloads with a cluster size which is ½ of earlier
~2x cost reduction with the same price performance

4TB TPCH	August: 10 HeatWave, 1 MDS node	Now: 5 HeatWave, 1 MDS node
Data per node	410G	820G
Performance	6.35 sec	11.59 sec
Annual cost	$34,073	$18,585
Price perf	$0.007	$0.007

- Intelligent pipelining and interleaving of decompression across cores, results in no overhead

Oracle also gives customers the option of pausing and resuming their MySQL databases without any data loss. For example, a development team can turn off their clusters for the night and resume at the same place in the morning. Hence, they only pay for the time that they are actually using the OCI resources.

Conclusion

One of the fabulous things about today's cloud industry is that the wicked competition among some of the world's top technology companies—Oracle, Microsoft, Google, Amazon, Snowflake, and others—drives a never-ending stream of innovation and value for customers. And Oracle's huge investments in and plans for MySQL HeatWave are a perfect example of that dynamic, as Larry Ellison and

company look to grab new share not by offering the same old stuff at 3% less than everybody else, but rather by boldly changing the rules of the game and tipping them heavily in favor of customers.

I've been around the tech business for some time, and I don't recall seeing performance gaps as profound as the ones that Oracle has cited for its new MySQL HeatWave database. When technological breakthroughs of this magnitude happen—and please trust me when I say that breakthroughs of this scale do not happen often—the result is not just small-scale incremental moves by competitors trying to play catch-up in an outdated game offering outdated results. Instead, the competitors are forced to regroup and head back to the labs and attempt to compete against entirely new standards, new levels of performance, new levels of value, and ultimately new levels of customer expectations. With that in mind, the best way to close out this analysis is to reiterate the benchmark results that show, right here and right now, which cloud vendor is driving truly unique levels of price-performance and business value for enterprise customers:

- Oracle MySQL HeatWave versus AWS Redshift: Oracle is 4.8X better
- Oracle MySQL HeatWave versus Snowflake: Oracle is 14.4X better
- Oracle MySQL HeatWave versus Microsoft Synapse: Oracle is 14.9X better
- Oracle MySQL HeatWave versus Google Big Query: Oracle is 12.9X better
- Oracle MySQL HeatWave ML versus AWS Redshift ML is 25X faster
- Oracle MySQL HeatWave ML versus AWS Redshift ML is 1% of the cost

In a data-driven business world that is only just beginning to fully understand the power of data to drive innovation, growth, and digital-economy success, the astonishing performance gaps that Oracle has achieved with its new MySQL HeatWave database are going to be impossible for business leaders to overlook.

---END---