



Oracle MySQL HeatWave Inflection Point

by, David Floyer

September 1st, 2021

Wikibon believes MySQL Heatwave is an inflection point in cloud databases. HeatWave converges OLTP and OLAP into a single database and accelerates MySQL queries by a factor of up to 1,000 times. In addition, by leveraging integrated hardware and software, HeatWave eliminates the need for lengthy ETL processes. As a result, it improves the ability to deploy smarter transactional applications, automating more complex business processes.

As a result, enterprise IT early adopters can expect MySQL HeatWave to perform about seven times faster than AWS Redshift or Snowflake at a 2-5 times lower cost. HeatWave was measured as 17.5 times faster and less than half the cost of AWS Aurora, which is 42 times better price-performance.

Alibaba, AWS, Google, Microsoft, and Snowflake will need some time to rearchitect their MySQL offerings and their IaaS & PaaS platforms to run converged databases properly. In addition, most of these vendors will need to invest heavily in other converged database hardware and software.

Wikibon strongly recommends that enterprise IT departments set a three-year plan to eliminate separate OLAP databases and ETL from MySQL transactional databases and move to MySQL converged databases that support transactional and analytic workloads. MySQL HeatWave, the only platform that currently delivers these characteristics, is the obvious platform to evaluate now and for the next few years.

Premise

Wikibon believes MySQL HeatWave is an inflection point in cloud databases. HeatWave converges OLTP and OLAP into a single database and accelerates MySQL queries by a factor of up to 1,000 times. In addition, by leveraging integrated hardware and software, HeatWave eliminates the need for lengthy ETL processes. As a result, it improves the ability to deploy smarter transactional applications, automating more complex business processes.

The service is an integrated OLTP/OLAP database. Additional data types and databases are likely to be integrated. As a result, ETL can become a distant memory for most small and midsize organizations. They can compete on equal terms with large organizations on the ability to automate and time to adapt.

Research Highlights

This Wikibon research addresses three central findings:

1. First, Oracle's database innovation capability offers users more than an order of magnitude improvement in price-performance by using MySQL HeatWave for analytic workloads compared to alternative cloud vendor offerings for analytic services.
2. The most significant business benefit is the potential to simplify IT by eliminating ETL and a second analytical database. ETL is the lengthy process of Extracting, Transforming, and Loading the data from Transactional Databases (OLTP) to a Data Warehouse database (OLAP). Managing one database is less than half the cost of managing two.
3. Cloud-based database services make it easier to add automation and optimization services. For example, Oracle introduced MySQL Autopilot, which makes extensive use of machine learning to automate provisioning, query execution, data loading, and failure handling. These automation and optimization services significantly improve programmer and analyst productivity.
4. The fourth finding is that developers can now combine OLTP & OLAP databases and integrate real-time analytics into a smarter transactional workflow. Lines-of-business can then improve business productivity by automating more complex business processes. In addition, MySQL HeatWave will allow ISVs with transactional MySQL systems to improve the functionality of their products by incorporating real-time analytics.

The following research evaluates these findings and the business implications for IT, users, and vendors.

Traditional MySQL Model & ETL

Oracle MySQL HeatWave Inflection Point

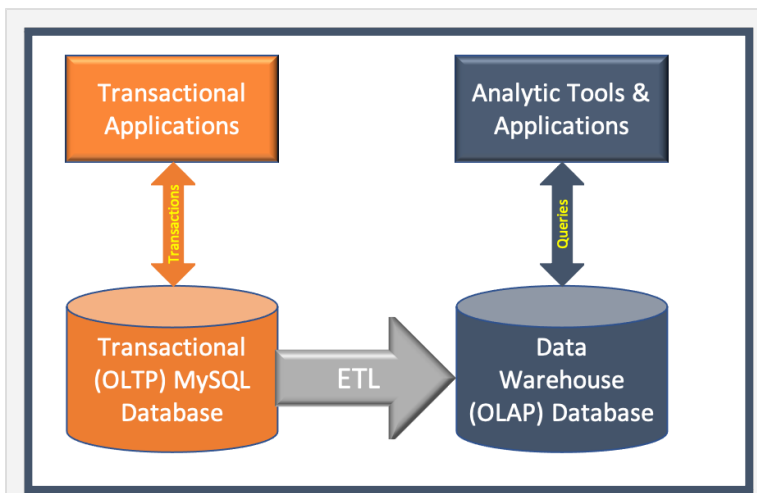


Figure 1 - Traditional MySQL Database Model, with Transactional MySQL Database feeding the Data Warehouse via ETL Processes.

Source: © Wikibon 2021.

Figure 1 illustrates the traditional MySQL deployment model. The left-hand side shows the transaction applications running against the MySQL database. The database is traditionally row-major, with quick point query database lookups.

MySQL has been traditionally slow for analytic processing, which is usually column-based. So instead, IT usually runs the analytic workloads on a specialized OLAP database, which speeds up ad-hoc and reporting analytics.

The downside of this traditional model is that IT must design and run ETL (Extract, Transform, Load) processes to take the data from the MySQL transaction database to the OLAP database. In addition to the cost of the OLAP database, the ETL software is expensive, inflexible, and temperamental, and the transfer time can take hours, days, or even weeks. CIOs are rarely happy with their ETL process. It is usually a significant source of data warehouse frustration for both IT and the lines of business, as data is stale by the time they analyze it.

Oracle MySQL HeatWave announcements

Oracle owns Open-source MySQL as part of the Sun acquisition completed in 2009. Enterprises have historically used MySQL for transaction processing applications, which are primarily row-based. As a result, many enterprise IT departments use MySQL for OLTP systems and then ETL the data to a second specialized OLAP database.

Oracle made a series of MySQL cloud announcements over the last year, radically improving MySQL query processing. These announcements centered around HeatWave, an in-memory query acceleration engine for MySQL Database Service in Oracle Cloud Infrastructure (OCI).

MySQL HeatWave has three major improvement components:

1. First, the HeatWave engineering team developed new algorithms for efficient distributed query processing, designed for massive scale. For example, the HeatWave engine has a highly partitioned architecture which enables a high degree of query parallelism. These advanced algorithms and techniques in HeatWave make query processing fast, scalable, and efficient.
2. HeatWave has been optimized for low-cost commodity cloud database services, including object storage, optimized compute, and commodity networking. This optimization reduces the cost of the service.
3. The recent announcement of MySQL Autopilot introduces an extensive use of machine learning to automate various aspects of the MySQL HeatWave service, including provisioning, data loading, and query processing. In addition, Autopilot efficiently collects data from query compilation and execution. This data allows dynamic choices for the execution plan, such as data placement and size of the partitions.

Oracle MySQL HeatWave Inflection Point

The expected improvements from MySQL HeatWave are faster query response, significant reductions in the resources required to run analytics, and reduced costs by eliminating the need for a separate OLAP database and ETL.

The following section measures the relative price-performance for HeatWave compared to cloud database alternatives. A later section entitled “Converged MySQL Model without ETL” looks at simplifying business processes by eliminating a separate OLAP database and ETL.

Oracle MySQL HeatWave Price-Performance Benchmarks

Overall Benchmark Results

Figure 2 below summarizes the benchmarks run by Oracle and others.

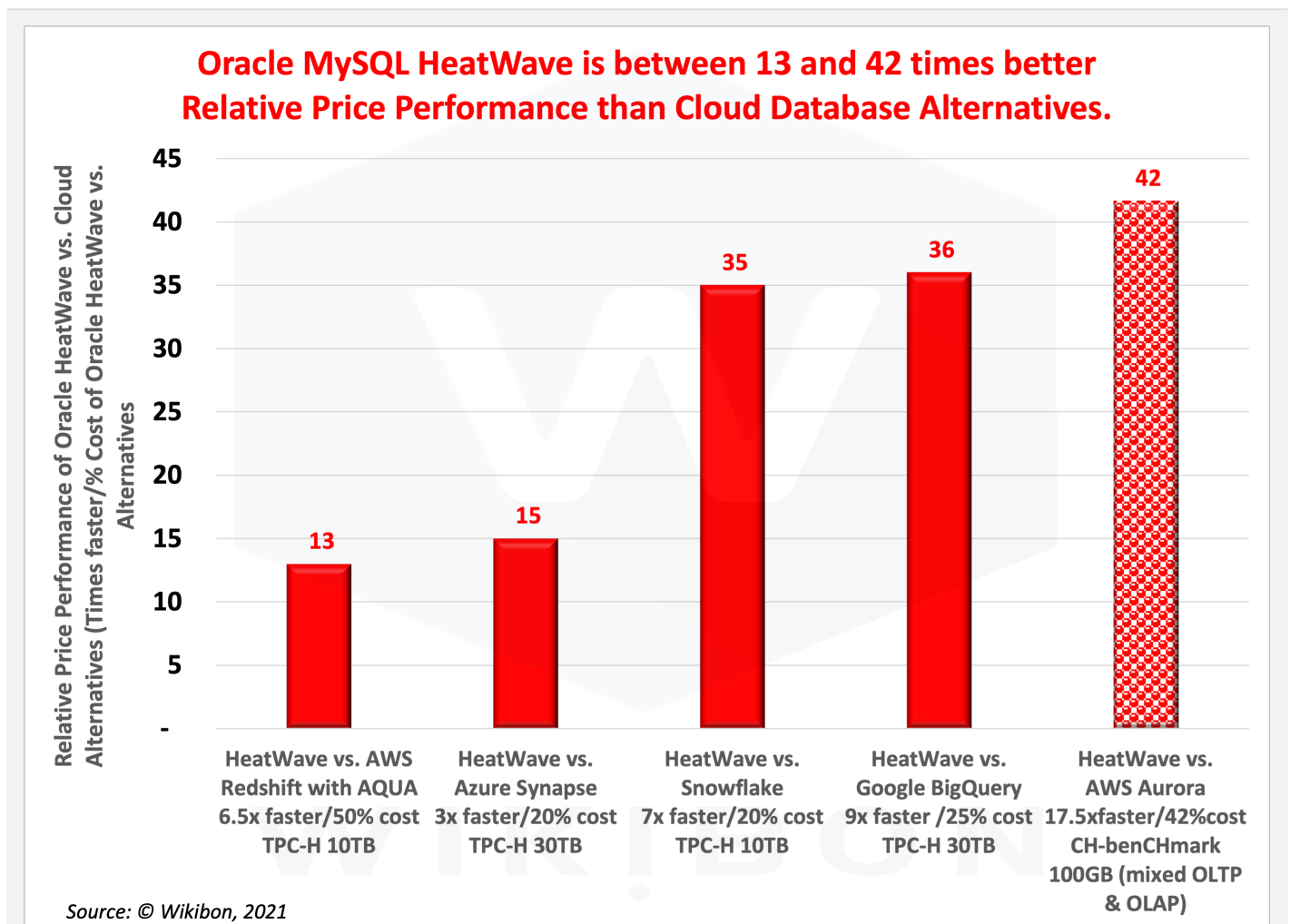


Figure 1 - Relative Price-Performance of Oracle MySQL HeatWave vs. Cloud Database Alternatives.

Data Source: Oracle 2021. Graphic Source: © Wikibon 2021.

The left-hand Y-axis in Figure 2 above shows the relative price-performance results of benchmark suites run on MySQL HeatWave vs. other cloud offerings. Note that this is a “Times” scale. For example, 1 on this scale represents price-performance equality; 13 means that HeatWave is thirteen times better price-performance than AWS Redshift with AQUA, and so on.

The x-Axis shows a comparison of HeatWave with 5 different services. The format of the titles at the bottom

of Figure 2 is the same for all five columns. As an example:

The first two lines show the alternative cloud database compared. For example, in the first column, the comparison is HeatWave vs. AWS Redshift with AQUA.

The height of the column shows the number of times greater HeatWave is compared to the alternative. For example, the first column shows the relative price-performance of HeatWave is 13 times better than AWS Redshift with AQUA.

The third line shows the calculation of relative price-performance. In the first column, HeatWave is 6.5 times faster than AWS Redshift with AQUA, at 50% of the cost. Thus, the price-performance calculation is **$6.5 \div 50\% = 13$** .

The last line is the benchmark used. For example, in the first column, the benchmark is the OLAP TCP-H benchmark with 10 Terabytes of Data.

TPC-H Derived Benchmarks

The first four columns in Figure 2 are OLAP benchmarks derived from TPC-H with either 10 or 30 Terabytes of data. They are in plain red. The benchmark uses MySQL HeatWave for the Oracle TCP-H benchmarks, and the Alternative OLAP runs the same TPC-H benchmark.

The TPC-H comparison metric is the geometric mean of all of the query response times. The price component of price-performance is calculated from the one-year list prices for all vendors except Snowflake.

Derived CH Benchmark

The 5th column is stippled red and gives the result of a different benchmark called “[CH-benchmark 100GB](#)”, which consists of a mixed workload based on TPC-C and TCP-H. The metric used to compare performance is the total elapsed time to complete the benchmark.

The two workloads compared in column 5 are HeatWave vs. AWS Aurora using MySQL as a base. Aurora has a parallel query option, which improves query performance by up to 5 times, with 2 times being a reasonable estimate of general query performance improvement.

The result of this benchmark is stunning. AWS Aurora takes 17.5 times longer to complete the CH benchmark, at 42% of the cost, even with the parallel query feature. The overall price-performance, calculated as HeatWave 17.5x faster at 42%cost, so HeatWave is 42x better.

Wikibon Benchmark Assessment

Benchmarks are data points. As always, real-world enterprise query workloads will differ from the TPC-H benchmarks. Likewise, mixed workloads will vary from the CH benchmark. Wikibon has often criticized Oracle’s benchmarks results.

However, in this instance, Wikibon assesses Oracle’s benchmark work as fair and reasonable. The benchmark parameters and scripts are all available on Github. Wikibon recommends enterprise specialists take a closer look at HeatWave’s performance and price-performance benefits and determine the relevance to their own workloads.

Oracle customer feedback from early deployments has reflected the same or better performance improvements in real-world environments.

Bottom Line: Price-performance for MySQL HeatWave:

Oracle has employed its database design and architectural capabilities to improve MySQL price-

Oracle MySQL HeatWave Inflection Point

performance dramatically. As a result, Wikibon assesses that enterprise IT should expect improvements in relative price-performance of between 13x and 42x using MySQL HeatWave compared to the best OLAP databases from AWS, Google, Microsoft, and Snowflake.

HeatWave 13x better than AWS Redshift with AQUA (6.5x faster/50% cost = 13x)

HeatWave 15x better than Azure Synapse (3x faster/20% cost = 15x)

HeatWave 35x better than Snowflake (7x faster/20% cost = 35x)

HeatWave 36x better than Google BigQuery (9x faster/25% cost = 36x)

HeatWave 42x better than AWS Aurora (17.5x faster/42% cost = 42x)

The next section looks at the even greater potential benefits of eliminating a separate OLAP database and ETL.

Converged MySQL Model Eliminates a Second Database and ETL

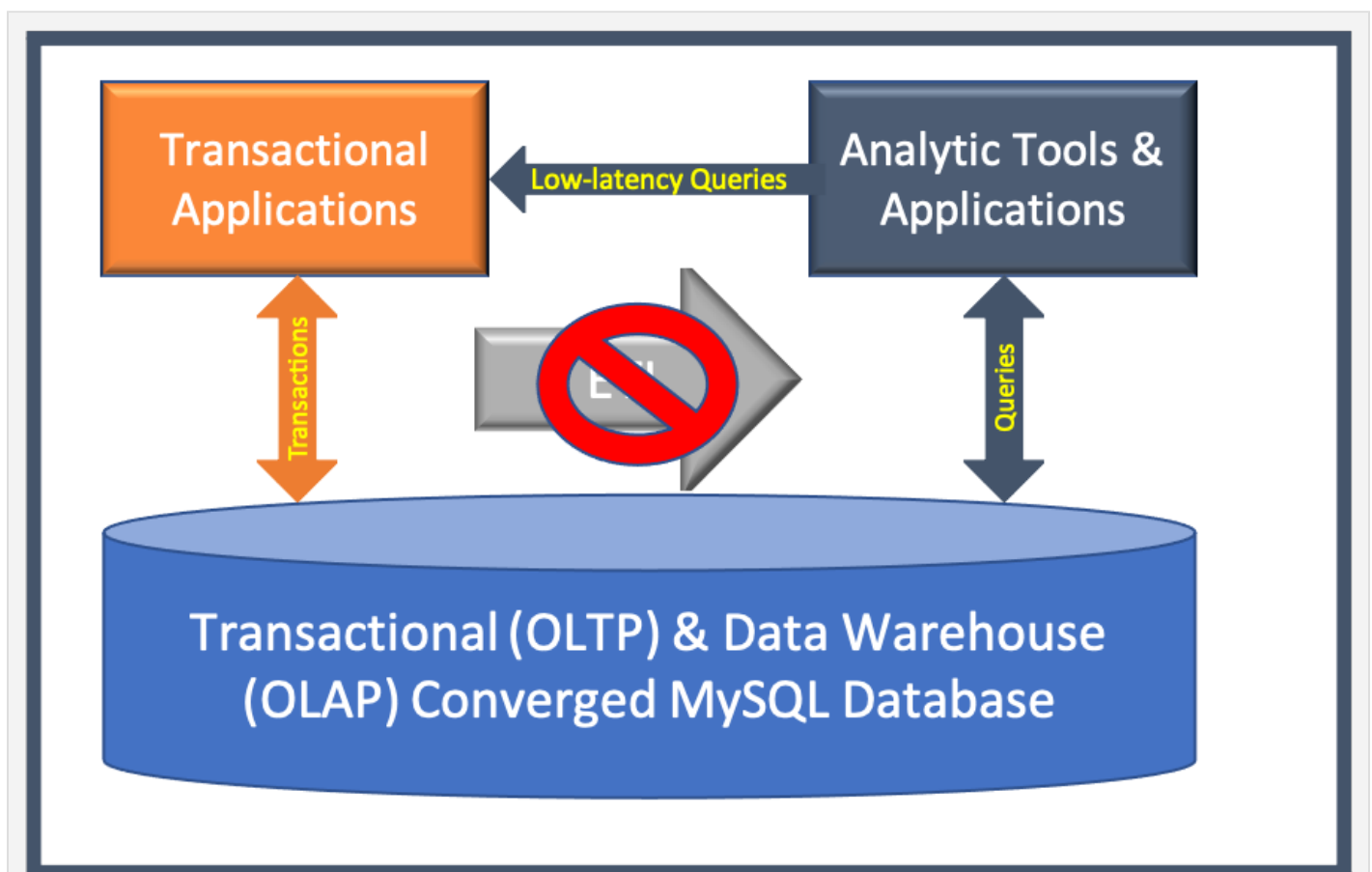


Figure 3 - Converged OLTP & OLAP MySQL database, eliminating the need for a separate OLAP database and ETL.

Source: © Wikibon 2021.

Figure 3 shows what is possible with the OLAP performance improvements from MySQL HeatWave.

To achieve this combined OLTP/OLAP database,

Transactions and Analytical queries run against the same converged MySQL database, as shown in blue

Oracle MySQL HeatWave Inflection Point

at the bottom of Figure 3. Transactional systems will use the row-based MySQL database schema, and the analytic systems will use the column-based extensions of HeatWave.

The converged MySQL database eliminates the need for a separate OLAP database and ETL.

The analytic system can now execute real-time low-latency queries initiated by the transactional applications and support ad-hoc and reporting queries from the analytical applications. The arrows illustrate this in Figure 3.

The system must be dynamically tuned for all workload types. Oracle has introduced MySQL Autopilot's Autoscheduling, prioritizing transactions and shorter queries over long-running queries to optimize overall performance.

Bottom Line: Wikibon believes that converged cloud databases allow dramatic improvements in database effectiveness. The elimination of a second OLAP database and ETL is a significant step in any digital transformation. It enables smarter transactional applications with far faster access to in-line real-time data. In turn, application designers can use real-time RPA tools to automate synchronous business processes.

Note: the cost for purchasing and managing a separate analytical database and ETL costs are not included in the price used in the price-performance calculations shown above. These costs would make the price-performance differentials between HeatWave and competitive offerings even greater.

MySQL Marketplace Implications

Introduction

Wikibon believes that the technology underlying MySQL HeatWave technology is an inflection point in database design and architecture. The parallelism and speed-up of the MySQL query handling enable IT to fuse MySQL transactional and analytic databases. Users can eliminate the ETL shown in Figure 1. The low cost of the combined package makes this an IT volume market. As a result, an accelerated learning curve will drive rapid future innovations and cost reductions.

[MariaDB](#) and Couchbase have started on the MySQL architectural road of combining OLTP and OLAP. In the high-end database market, Microsoft, Oracle, and SAP HANA offer converged databases. However, combining the innovative architecture and cloud to improve function and lower costs makes HeatWave a unique offering for users. Of course, any migration is hard work, but the financial payback time is quick and the strategic benefits high.

The following sections are a brief analysis of the impact of HeatWave on the leading database vendors.

AWS Database

AWS has a successful and varied portfolio of specialized databases. AWS' strategy is to take open-source code and engineer improvements to increase performance on AWS cloud infrastructure. AWS claims this gives it flexibility and time-to-market advantage for the different types of databases. In addition, AWS has taken a "right database for the job" approach and provides access to sixteen different database types. These include Aurora (mainly transactional) and Redshift (analytic).

AWS's priority is to provide low-cost processors, storage, and networking resources. As a result, AWS has successfully virtualized commodity hardware and can sell the same resources simultaneously to many customers and reduce costs for themselves and their customers. This approach works well for smaller and less complex databases, which is the majority by volume of the database marketplace.

However, Figure 2 above shows that HeatWave has 42 times better price performance in a mixed OLTP/OLAP environment than AWS Aurora, with no ETL cost, no second database to manage, and superior

Oracle MySQL HeatWave Inflection Point

functionality for MySQL developers. 42x comes from 17.5 better performance and 42% lower cost – this is a huge difference! We believe customers will demand that AWS provide the similar converged functionality as HeatWave, which will be attractive to the MySQL sweet spot, small and mid-sized companies, and independent departments of large enterprises. MySQL users represent over 50% of the database market.

Wikibon is not aware of any MySQL open-source projects which could be used as a basis for an AWS offering. This means that AWS will need to develop this functionality themselves. AWS has started the approach with the Parallel Option for Aurora and will need to take it much further in our view and integrate other databases and data types.

Wikibon believes that AWS will respond quickly behind the curtain. AWS will also want to retain MySQL users. New entrants such as Snowflake will need improved cloud technology fast to stay competitive with HeatWave and others (see Snowflake section below).

Google BigQuery

Google is in the worst TPC-H benchmark position in Figure 2 with a 9x performance deficit and 4x more expensive.

This is not a surprise, as the TCP-H workload is not in the sweet spot of Google BigQuery. Its strength is its ability to handle enormous amounts of data quickly. Data scientists are drawn to its eclectic approach to data mining, and BigQuery fits well into machine learning.

Many enterprise workloads, especially those in the MySQL space, are simpler. Previously, traditional solutions such as AWS Redshift were superior in handling normal business processes and schemas that are not too complex.

The advent of MySQL HeatWave has made it harder for Google BigQuery to compete in the general-purpose data warehouse market. However, Wikibon believes HeatWave mandates Google take a similar approach within a few years.

Microsoft

Microsoft Azure offers the best MySQL performance TPC-H benchmark position relative to the other cloud database providers shown in Figure 2. The performance deficit compared to HeatWave is 3x. Also, Wikibon believes Microsoft has the database software and hardware design skills to lower that deficit and make its service more cost-competitive. Wikibon believes that Microsoft will execute the same product strategy as Oracle HeatWave to compete against AWS and get there first.

Oracle

Oracle has a unique hybrid database cloud strategy. It offers a converged Oracle Database based on the same software and hardware delivered on the Exadata platform, traditional on-premises, Cloud@Customer, and the OCI cloud. Larger Oracle customers have experienced significant difficulties in moving Oracle enterprise workloads to AWS. Wikibon believes that Oracle will hit escape velocity with its OCI Autonomous database products.

However, Oracle faces a MySQL Heatwave marketing challenge in dealing with the “if it ain’t broke, don’t fix it” market mindset. Large customers will continue to migrate large MySQL systems, which have hit performance bottlenecks, to HeatWave. However, the MySQL market is huge in number, incredibly diverse, and international. Over 50% of database users deploy MySQL, including giants such as Facebook. Support is often a local wiz-kid.

Oracle’s challenge is building a MySQL open-source culture and a way of appealing to this vast user and supporter ecosystem. Oracle must move fast. The MySQL HeatWave technology is by far the best in the market now, but vendors can copy ideas. Is licensing the technology the way forward? Is attracting startups

such as Couchbase and Snowflake a way forward? How can Oracle attract ISVs? It will be fascinating to see Oracle moving forward with this product and avoiding infanticide from its bigger brother. On balance, it is a great problem to have.

Snowflake

Snowflake has created significant market momentum and focuses on reducing complexity with great ease-of-use and time-to-value for data warehouses. Snowflake has also introduced enhanced capabilities for sharing data warehouses. Wikibon believes in the importance of making data discoverable and shareable so enterprises can build data products and data services in a securely governed environment.

However, just like AWS, Snowflake requires customers to ETL data from their transactional databases into the Snowflake cloud, even if they make it easier. We believe over time, Snowflake will need to integrate transactional MySQL products to enable access to in-line real-time data as with HeatWave.

More serious for Snowflake's business model is its dependence on cloud vendors to provide infrastructure services for its database, for which it pays directly. Snowflake secured extremely favorable prices with AWS, its primary supplier. As for performance and price-performance, the benchmark results in Figure 2 show that Snowflake running on AWS is seven times slower and five times more expensive than HeatWave. In other words, HeatWave's price-performance is 35x better than Snowflake.

However, clearly, Snowflake has customers who will pay a premium for the ease-of-use features and the ability to move work out to the line of business. Wikibon believes such a high-cost difference for long-running queries will mean enterprise movement to MySQL HeatWave! The price difference will attract management's attention.

That leads to a dilemma for Snowflake. It can hope that AWS will improve its database infrastructure functionality or invest in its own hardware platform (which it says it is not doing). [Andreessen Horowitz has some interesting observations on this dilemma](#). The analysts in the research referenced in the previous sentence argue that SaaS companies are financially crazy not to start in the cloud but also financially crazy to depend on third-party cloud providers after they reach a certain revenue level. Never say never in the cloud!

Maybe Snowflake would be interested in using Oracle technology? Stranger bedfellows exist!

Conclusions & Recommendations

Research Findings

At the beginning of this research, we laid out three findings.

1. Oracle's database innovation capability offers MySQL users more than an order of magnitude improvement in price-performance by using MySQL HeatWave for query workloads compared to alternative cloud vendors' offerings.

Wikibon concludes that users will achieve about two orders of magnitude improvements in price-performance for MySQL query workloads.

Figure 2 shows clearly that MySQL HeatWave will be over an order of magnitude better price-performance than other cloud alternatives.

Wikibon believes that this competitive advantage is sustainable for at least three years.

Snowflake has ease-of-use benefits and some capabilities for sharing data warehouses, with potential Data Mesh capabilities. However, Wikibon would recommend that if line-of-business wants to pay Snowflake's high prices for ease-of-use, IT should identify long-running reporting queries and migrate

them to MySQL HeatWave or similar services.

2. A significant benefit for all businesses is the potential to simplify IT by eliminating a separate OLAP database and ETL. ETL is the lengthy process of Extracting, Transforming, and Loading the data from Transactional Databases (OLTP) to the Data Warehouse Databases (OLAP).

The benefits of eliminating ETL are in addition to the cost and performance benefits of using converged MySQL HeatWave.

Eliminating separate analytic databases and the ETL software reduces IT costs and reduces software management costs. Wikibon's experience leads us to believe that the savings from eliminating a second database and ETL will be significantly greater than the savings in section 1 above.

Note that the cost for purchasing and managing a separate analytical database and ETL costs are not included in the cost analysis used in the earlier price-performance calculations. Running two databases is greater than twice the cost of running one.

Eliminating ETL means the user analytic data is never out of sync with the transactional data.

Synchronous data allows lines-of-business better control of their business and is an essential step to Digital Transformation.

MySQL Autopilot makes extensive use of machine learning to automate provisioning, query execution, data loading, and failure handling. This will continue to improve the productivity of developers and systems administrators.

Wikibon recommends that MySQL users in small and mid-sized enterprises and independent departments in large enterprises plan to eliminate ETL within three years.

3. The third finding is that developers can now combine OLTP & OLAP databases and integrate in-line real-time analytics as part of a smarter transactional workflow. Lines-of-business can then radically improve business productivity by automating more complex business processes.

The potential benefits of using in-line real-time analytics to improve business processes and interactions between company staff, partners, suppliers, and customers can lead to full automation of business processes. The cost savings of a fully automated business process are usually over 30% direct to the bottom line, with improved customer and supplier satisfaction.

The ease of use for users and ISV developers using HeatWave will improve the application functionality and time-to-market.

Summary

Wikibon believes that MySQL HeatWave has delivered an important capability to MySQL users. As a result, enterprise IT early adopters can expect MySQL HeatWave to perform about seven times faster than AWS Redshift or Snowflake at a 2-5 times lower cost. AWS Aurora users should expect higher benefits.

However, Wikibon believes the most important user benefits of converged OLTP and OLAP databases will come longer term from eliminating ETL and developing smarter transactional applications to automate complex business processes.

Alibaba, AWS, Google, Microsoft, and Snowflake will need some time to rearchitect their IaaS & PaaS platforms to run converged databases properly. In addition, most of these vendors will need to invest heavily in other converged database hardware and software.

Enterprises of every size should plan to run a converged MySQL database for OLTP & OLAP and set a date to eliminate ETL within four years. In addition, users should expect and encourage transactional ISVs to

provide integrated OLTP/OLAP databases in their products. As a result, lines of business can plan to integrate in-line real-time analytics with systems of record and radically improve business process automation and costs.

Action Item

Wikibon strongly recommends that enterprise IT departments set a three-year plan to eliminate separate OLAP databases and ETL from MySQL transactional databases and move to MySQL converged databases that support transactional and analytic workloads. MySQL HeatWave is the only platform that delivers these characteristics and is the obvious platform to evaluate now and for the next few years.

David Floyer spent more than 20 years at IBM, holding positions in research, sales, marketing, systems analysis and running IT operations for IBM France. He worked directly with IBM's largest European customers, including BMW, Credit Suisse, Deutsche Bank and Lloyd's Bank. Floyer was a Research Vice President at International Data Corporation (IDC) and is a recognized expert in IT strategy, economic value justification, systems architecture, performance, clustering and systems software.



David Floyer

@dfloyer

david.floyer@wikibon.org