

ORACLE

MySQL HeatWave on AWS

One MySQL Database for OLTP, OLAP, and Machine Learning on AWS

Copyright © 2022, Oracle and/or its affiliates
Public

Purpose statement

This document provides an overview of features and enhancements included in HeatWave. It is intended solely to help you assess the benefits of HeatWave and to plan your I.T. projects.

Disclaimer

This document in any form, software or printed matter, contains proprietary information that is the exclusive property of Oracle. Your access to and use of this confidential material is subject to the terms and conditions of your Oracle software license and service agreement, which has been executed and with which you agree to comply. This document and information contained herein may not be disclosed, copied, reproduced or distributed to anyone outside Oracle without prior written consent of Oracle. This document is not part of your license agreement nor can it be incorporated into any contractual agreement with Oracle or its subsidiaries or affiliates.

This document is for informational purposes only and is intended solely to assist you in planning for the implementation and upgrade of the product features described. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described in this document remains at the sole discretion of Oracle. Due to the nature of the product architecture, it may not be possible to safely include all features described in this document without risking significant destabilization of the code.

Benchmark queries are derived from the TPC-H and TPC-DS benchmark, but results are not comparable to published TPC-H and TPC-DS benchmark results since they do not comply with the TPC-H TPC-DS specification.

Table of contents

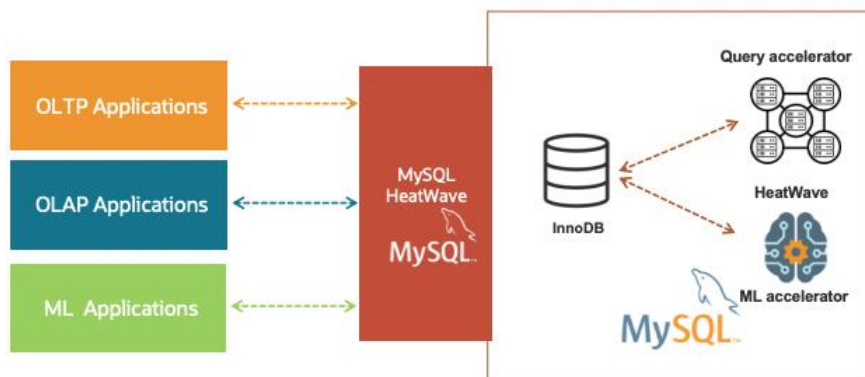
Purpose statement	2
Disclaimer	2
Table of contents	3
Introduction	4
Performance and Price Performance Advantages	5
Architecture of MySQL HeatWave on AWS	6
MySQL HeatWave Control Plane	6
MySQL HeatWave Data Plane	7
MySQL HeatWave Features on AWS	7
MySQL Database	8
HeatWave Analytical Processing	9
HeatWave AutoML	9
MySQL HeatWave Console	9
Lifecycle Management	10
Interactive Schema and Data Management	11
Interactive Query Interface & Workload Monitoring	12
Performance Monitoring	13
HeatWave AutoML Model Creation, Prediction, and Explanation	14
Conclusion	15

Introduction

MySQL HeatWave on Amazon Web Services (AWS) is a fully managed database service. The service automates the common management tasks of a database system, such as:

- Setting up, configuring and tuning the database
- Security patching and upgrading of the host operating system and database system
- Orchestrating database backups

MySQL HeatWave is the only service that combines transaction processing, real-time analytics, and machine learning within one single MySQL database.



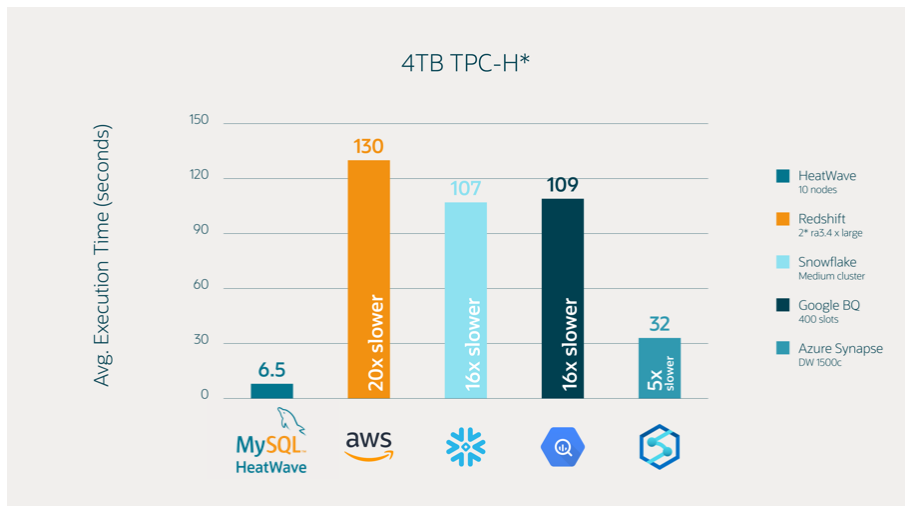
MySQL HeatWave eliminates the need for complex, time-consuming ETL operations and unnecessary data duplication between separate databases and tools for OLTP, analytics, and machine learning use cases. Customers avoid the latency and security risks of data movement between data stores while reducing costs. MySQL HeatWave also includes MySQL Autopilot, providing workload-aware, machine learning-powered automation of various aspects of the database system lifecycle; including provisioning, data management, query execution and failure handling for both OLTP and analytics workloads.

Oracle now makes all these MySQL HeatWave capabilities, which are built, managed, and continuously supported by the MySQL HeatWave development team, available on AWS. All components of the MySQL HeatWave service on AWS, namely the service console, control plane, and data plane, are built and optimized for AWS. MySQL HeatWave's native integration with AWS enables customers with applications already deployed in AWS to benefit from MySQL HeatWave without incurring the latency associated with accessing a database service running outside of AWS. Customers also don't incur the high data egress fees charged by AWS that would be necessary to migrate data to a service running outside of AWS. Lastly, the tight integration of MySQL HeatWave with the AWS infrastructure, e.g. Amazon S3, CloudWatch, PrivateLink, makes it easy for developers to rely on MySQL HeatWave for new applications.

“MySQL HeatWave on AWS fits perfectly into our data platform with 60X to 90X faster complex queries compared to AWS RDS and Aurora. It generates real-time analytics we need for targeted, multichannel campaigns. We now have greater scalability to onboard more data and new clients of any size without increasing IT admin.”

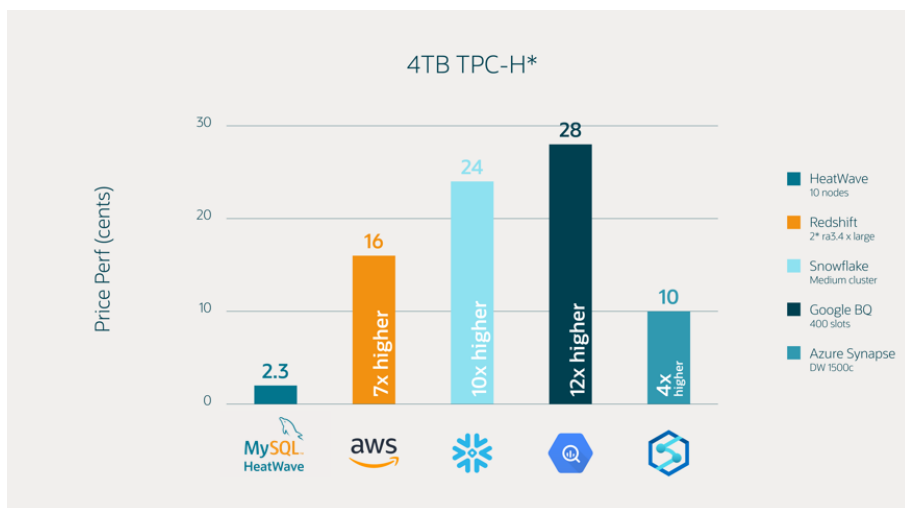
Thomas Henz
Chief Executive Officer, Johnny Bytes

Performance and Price Performance Advantages



“MySQL HeatWave on AWS has 50X faster complex queries compared to AWS RDS that provide us real-time insights to accelerate application development and help us improve patients’ lives.”

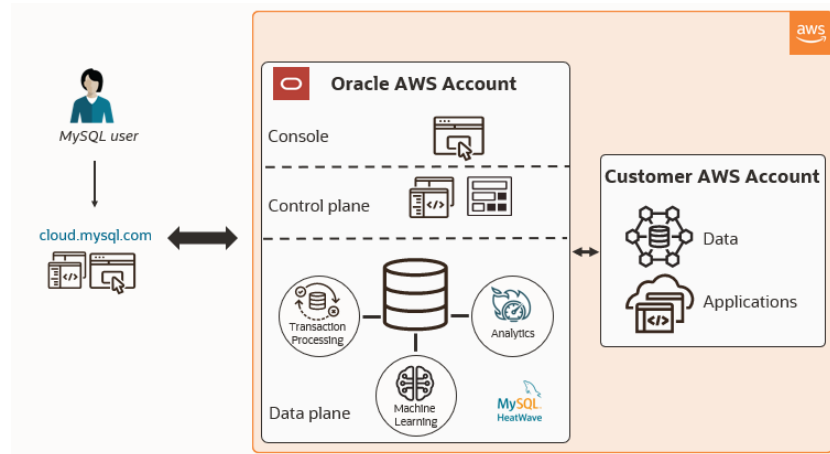
Kyle Yang
Assistant Manager
Bionime



*Only compute costs are considered in the above graphs. Pricing for Redshift is based on 1 year reserved instance, paid upfront. Pricing for Snowflake is based on standard edition. Pricing for Google Big Query is based on monthly flat rate commitment. Pricing for Azure Synapse is based on 1 year reserved pricing.*Benchmark queries are derived from the TPC-H benchmarks, but results are not comparable to published TPC-H benchmark results since these do not comply with the TPC-H specifications.*

With its superior data-processing architecture and optimization for AWS infrastructure, MySQL HeatWave on AWS delivers unmatched performance and price-performance. On the 4TB TPC-H benchmark, MySQL HeatWave on AWS delivers 7x better price performance than Amazon Redshift, 10x better than Snowflake, 12x better than Google BigQuery and 4x better than Azure Synapse. For machine learning, MySQL HeatWave on AWS is 25X faster than Redshift ML. On a 10GB TPC-C workload, MySQL HeatWave offers up to 10X higher and sustained throughput compared to Amazon Aurora at high concurrency. All of these fully transparent benchmark scripts are available on [GitHub](#) for customers to replicate.

Architecture of MySQL HeatWave on AWS



“MySQL HeatWave on AWS has 139X faster complex queries in comparison to AWS RDS and Aurora. This provides a significant opportunity to simplify the existing data infrastructure for both OLTP and OLAP, along with sub-second response time, to deliver an optimal experience.”

Anish Kumar
Associate Vice President
6D Technologies

MySQL HeatWave on AWS delivers a true native experience for AWS customers. The console, control plane, and data plane completely reside in AWS and are responsible for managing the MySQL HeatWave database resources in AWS. The control plane communicates with Oracle Cloud Infrastructure (OCI) Identity for account management, and with OCI metering & billing for monitoring and managing the usage and expenses associated with the customer’s account.

Once the user signs up for an OCI cloud account and registers their OCI account with MySQL HeatWave on AWS, the main interactions with the MySQL HeatWave service take place in AWS, through the service console hosted at cloud.mysql.com. The MySQL HeatWave console relies on the RESTful API provided by the MySQL HeatWave control plane to handle the user requests.

MySQL HeatWave Control Plane

The MySQL HeatWave control plane, which enables the management of MySQL HeatWave and maintains the necessary metadata, is built on publicly available AWS services. The control plane components are hosted in the Oracle AWS account dedicated to the MySQL HeatWave service and are tightly controlled with AWS Identity and Access Management permissions and policies.

The MySQL HeatWave control plane is responsible for the management of the database system lifecycle including provisioning/de-provisioning/pausing, the configuration of the database system, orchestrating backups, security patching, upgrades, monitoring, as well as ensuring the isolation of different database systems from each other.

MySQL HeatWave Data Plane

MySQL HeatWave on AWS hosts all of the customer databases components in a dedicated AWS account and strictly isolates them from the service control plane components and other database systems managed by the control plane.

MySQL HeatWave databases are hosted on AWS services, which are publicly available to AWS customers, such as Elastic Compute Cloud (EC2) for data processing, Elastic Block Storage (EBS) for storage, and VPC for resource isolation. Customers do not have access to the infrastructure on which the database runs, such as the database host machines or storage; they are provided with a MySQL endpoint, which can be accessed by a standard MySQL client. Customers can restrict the client addresses that can access their database systems through the MySQL endpoint and the connection between the client and database system is secured with TLSv1.2. Using the MySQL endpoint, customers can run any application against their MySQL HeatWave database systems.

MySQL HeatWave Features on AWS

MySQL HeatWave provides a single database system for high-performance and secure transaction processing, real-time analytics, and machine learning. Each MySQL HeatWave database system consists of a MySQL Database node, mainly targeting transaction processing, and a configurable number of HeatWave nodes for analytical processing and machine learning tasks. MySQL HeatWave on AWS strives to provide the best price/performance by optimizing all components of the underlying AWS infrastructure—while keeping costs as low as possible.

The built-in MySQL Autopilot provides workload-aware, machine learning-based automation of various aspects of the application lifecycle, including provisioning, data management, query execution, and failure handling. MySQL Autopilot features include auto provisioning, auto parallel loading, auto encoding, auto data placement, auto scheduling, auto query plan improvement, auto change propagation, and auto error handling. Combined, these features improve the performance of the application, reduce cost by predicting the optimal configuration to run a workload, and reduce manual database administration. MySQL Autopilot's machine learning models, hosted in the database system, leverage the run time information like data and query statistics, memory, and network usage, providing continuous improvements over time. In addition to the automation features offered for the HeatWave analytics engine, MySQL Autopilot includes capabilities designed for OLTP workloads, which further improve the MySQL HeatWave price performance (more information provided further below).

Oracle provides the latest MySQL Enterprise Edition version in MySQL HeatWave on AWS, which contains the state-of-the-art performance, functionalities and security features developed, maintained, and supported by the MySQL team at Oracle. With each new release of the MySQL server, customers of MySQL

“HeatWave with MySQL Autopilot on AWS is a gift from the database gods.”

Matt Kimball
Senior Analyst
Moor Insights & Strategy

HeatWave on AWS will have immediate access to the new features that are developed and tailored for the AWS infrastructure to maximize performance and security. Some of the performance and security features in the first release of MySQL HeatWave on AWS are presented in the following sections.

MySQL Database

- **Tuned for peak performance:** MySQL HeatWave on AWS provides several shapes (i.e., AWS EC2 instance types) with different vCPU counts and memory sizes, on which the MySQL server can run for transaction processing. Customers can select a shape based on their workload needs. Each MySQL shape comes with a default MySQL configuration tailored for that shape to maximize performance on the target AWS infrastructure. Depending on the selected shape and storage size specified by the customer, the underlying storage system built on AWS EBS is tuned for peak performance without magnifying the associated costs.
- **New MySQL Autopilot features for OLTP:** MySQL Autopilot is now enhanced with two new features to provide workload-aware, machine learning-powered automation capabilities for OLTP workloads—to further improve performance and throughput while enabling cost savings:
 - o **Auto Shape Prediction:** To alleviate the burden of experimenting with different MySQL shapes to determine the most performant shape for a given workload, Auto Shape Prediction provides suggestions for the right MySQL server shape, based on highly accurate predictions from machine-learning models inside the MySQL server and the most recent query execution metrics and traces. Since Auto Shape Prediction continuously collects workload execution statistics, it can adapt to the evolving workload patterns and hence provide the customer with suggestions based on the most recent workload.
 - o **Auto Thread Pooling:** With Auto Thread Pooling, MySQL HeatWave prioritizes not only peak single-thread performance, but also high throughput in the presence of concurrent clients running concurrent queries on a MySQL server. With this feature, the MySQL server now can perform workload-aware admission control of the incoming transactions. It eliminates the resource contention created by too many awaiting transactions, automatically queuing them to maximize performance while sustaining the throughput in the face of high concurrency.
- **Advanced security and compliance features:** MySQL HeatWave on AWS includes several comprehensive security features natively implemented in the MySQL server, as opposed to other services such as Amazon Aurora, which provide security methods as an additional layer on top of the database.
 - o **Data masking and de-identification:** Helps organizations protect sensitive data from unauthorized users by hiding and replacing real values with substitutes.

“For cost conscious IT teams and developers, MySQL HeatWave on AWS represents a whole new TCO calculation with zero cost for what are add-on services on AWS and no data egress fees.”

Marc Staimer
Senior Analyst
Wikibon

- **Asymmetric encryption:** Enables developers and DBAs to increase the protection of confidential data and comply with regulatory requirements including HIPAA, Sarbanes-Oxley, and the PCI Data Security Standard, through encryption, key generation, digital signatures, and other cryptographic features.
- **Database firewall:** Provides real-time protection against database-specific attacks such as SQL injections, by monitoring, alerting, and blocking unauthorized database activity without any changes in the application.

HeatWave Analytical Processing

- **Optimized and tuned for peak performance and best price performance:** As the MySQL server, the HeatWave analytical processing engine is tuned to achieve the best performance on the AWS infrastructure. To minimize costs on AWS, the HeatWave processing engine is enhanced with aggressive compression of the in-memory data, which helps to reduce the memory footprint, hence the number of required processing nodes, without sacrificing peak performance.
- **Support for smaller datasets:** HeatWave has been mainly designed to meet the needs of very large datasets by utilizing many processing nodes with large memory capacities. To enable OLTP databases with smaller datasets to benefit from the capabilities of the HeatWave engine, MySQL HeatWave on AWS provides customers the option of selecting a smaller HeatWave cluster shape, enabling them to benefit from the high performance of HeatWave at a lower cost.

HeatWave AutoML

HeatWave AutoML, integrated in the MySQL HeatWave service, provides fully automated in-database machine learning capabilities including training, inference, and explanation. This allows customers to securely leverage machine learning on real-time data without the complexity, latency, and cost of ETL. Since HeatWave AutoML stores all trained models inside the MySQL database, it eliminates the need for moving the models out of the MySQL HeatWave database system to an external machine learning tool or service.

In addition to enabling all the existing HeatWave AutoML features on AWS through the SQL interface, the MySQL HeatWave console on AWS enables HeatWave AutoML features through a user-friendly interface, further increasing the ease of use of the advanced HeatWave AutoML capabilities. To learn more about HeatWave AutoML, download our [technical brief](#).

MySQL HeatWave Console

The MySQL HeatWave console is designed to facilitate:

“For any developers working with MySQL on AWS, Oracle has just dropped a big productivity boost on your doorstep without the big price tag.”

Carl Olofson
Research Vice President
IDC

- Lifecycle management of the MySQL HeatWave resources, such as MySQL Database systems, MySQL backups, and HeatWave clusters.
- Resource and HeatWave data management by providing relevant MySQL Autopilot capabilities in the console.
- Workload monitoring through visualization of the schema metadata and statistics that are stored on the MySQL server.
- Management and visualization of training, inference, and explanation of machine learning algorithms offered by HeatWave AutoML.

Lifecycle Management

The screenshot displays the MySQL HeatWave console interface. At the top, there are tabs for 'DB Systems' and 'Backups'. Below the tabs, there is a 'Create MySQL DB System' button and several control buttons: 'Start', 'Stop', 'Restart', and 'Actions'. A search bar is also present with a dropdown for 'State' set to 'Any State' and a 'Search Name' field.

Name	State	HeatWave Cluster	HeatWave State	Created
name_710	Deleted	-	-	Sun, 19 Sep
name_665	Active	heatwave-cls-test-4	Active	Thu, 02 Sep
name_678	Active	heatwave-cls-test-3	Active	Fri, 13 Aug
name_511	Active	-	-	Thu, 15 Jul
name_489	Active	heatwave-cls-test-2	Inactive	Thu, 15 Jul
image-8.0.25-runs	Active	heatwave-cls-test-1	Active	Sat, 03 Jul

Below the table, there is a section for 'MySQL DB System Details' for the selected resource 'name_511'. This section is divided into 'Summary' and 'DB System Information'.

Summary:

- Name: name_511
- State: Active
- Resource ID: cee86d89-9254-4924-8680-5c8e3183bd51

DB System Information:

General Information	DB System Configuration
Description: description_511	Storage Size: 50 GiB
Created: Thu, 15 Jul 2021 13:43:07 GMT	MySQL Version: 8.0.27-u2-cloud

At the bottom of the console, there is a navigation bar with tabs for 'Introduction', 'MySQL', 'HeatWave Clusters', 'Workspaces', and 'Performance'.

Customers can manage the MySQL HeatWave resources associated with their accounts through the MySQL HeatWave console. The MySQL HeatWave console provides a single pane of glass to manage a MySQL Database system, the HeatWave cluster attached to the MySQL Database system, as well as the backups of the database system. The service console allows users to, for example, provision/deprovision/configure resources, pause a MySQL database system, and pause the HeatWave cluster attached to it if no user activity is expected in order to save costs.

Interactive Schema and Data Management

The screenshot shows the 'Estimate Cluster Size with MySQL Autopilot' interface. It features a 'Refresh Estimate' button and a timestamp 'Estimate last refreshed on Fri, 30 Sep 2022 17:42:36 GMT'. Below this, there are two main tables. The left table lists schemas, and the right table lists tables from selected schemas. At the bottom, there are four summary metrics: '183.470 (GiB) Memory required by the schemas/tables selected', '256 (GiB) Memory provided per node', '1 HeatWave Cluster nodes required', and '256 (GiB) Memory provided by 1 node cluster'. A 'Cancel' button and an 'Apply Cluster Size Estimate' button are also visible.

Name	HeatWave Cluster Memory Usage (GiB)	Tables Selected	Warnings
tpch_1024	183.470	8 of 9	1

Name	Warnings	Memory Size Estimate (GiB)	Rows Estimate
tpch_1024.CUSTOMER		8.673	145,758,420
tpch_1024.LINEITEM		683.208	5,843,973,733
tpch_1024.NATION		0.003	25
tpch_1024.ORDERS		119.521	1,458,277,490
tpch_1024.PART		17.992	196,258,911
tpch_1024.PARTSUPP		36.816	781,485,771
tpch_1024.REGION		0.003	5
tpch_1024.SUPPLIER		0.459	10,112,318
tpch_1024.tbl_with_unsupporte...		0.003	0

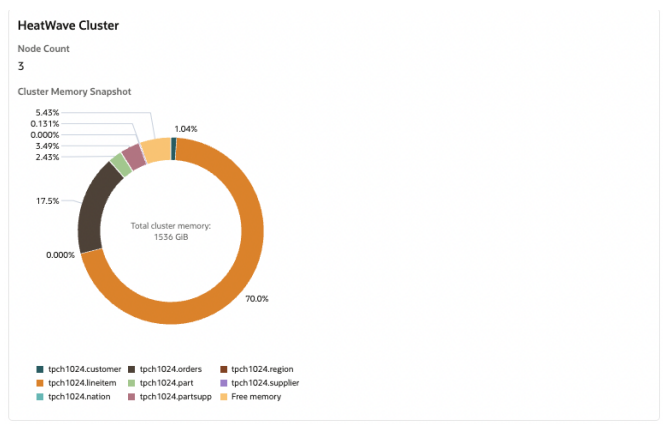
Once the customer provisions MySQL HeatWave and imports data into the MySQL server, they can leverage the Auto Provisioning feature of MySQL Autopilot to estimate the HeatWave cluster size needed for their workload. To enable this, the MySQL HeatWave console provides users with a list of schemas and tables in each schema in the MySQL server. Depending on which tables the user wants to run analytical queries on, Auto Provisioning estimates the total memory usage in the HeatWave cluster memory, and based on the HeatWave shape selected by the user, it estimates the number of HeatWave cluster nodes needed to accommodate the target dataset. Users then can provision a HeatWave cluster based on the number of nodes suggested by Auto Provisioning and load the desired tables to the HeatWave memory for analytical processing.

The screenshot shows the 'DB System' configuration page. It features a dropdown menu for the database system name, currently set to 'name_678', and a 'Disconnect' button. Below this, there is a section for 'Select tables to unload or load into HeatWave.' with a 'Detail View' toggle. A table lists the tables to be loaded or unloaded, with columns for Name, Memory Size Estimate, Rows Estimate, Load Status, String Column Encoding, and Predicted Load Time (s). The 'nation' table is selected for loading.

Name	Memory Size Estimate	Rows Estimate	Load Status	String Column Encoding	Predicted Load Time (s)
tpch1024	1454.08				
customer	16.04	138059807	100%	DICT: 3/8, VARLEN: 5/8	32.0
lineitem	1075.20	5756526822	100%	VARLEN: 16/16	343.0
nation	0.00	25	0%	VARLEN: 4/4	1.0

Customers can continuously monitor the schemas and tables loaded in the MySQL server as well as the HeatWave memory. This monitoring information helps loading the desired tables into the HeatWave cluster memory in the most

efficient way. The user first needs to select which tables to load into the HeatWave memory and then use Auto Parallel Load, which optimizes the load time and memory usage of the data load operation into HeatWave by predicting the optimal degree of parallelism for the set of tables selected by the user. To provide better visibility into storage and memory usage, the MySQL HeatWave console also provides detailed information about the estimated memory footprint of each table in the HeatWave cluster memory, encoding type, load status, as well as the predicted load time provided by MySQL Autopilot.



Once the desired tables are loaded into the HeatWave cluster memory, customers can monitor the actual breakdown of the memory consumption with the detailed breakdown of each table’s in-memory footprint.

Interactive Query Interface & Workload Monitoring

MySQL HeatWave ORACLE

Workspaces

DB System
Image: 8.0.25-nurs

Select tables to unload or load into HeatWave:

- tpch1024
 - customer
 - lineitem
 - nation
 - orders

HeatWave Cluster
Node Count: 3

Query 1

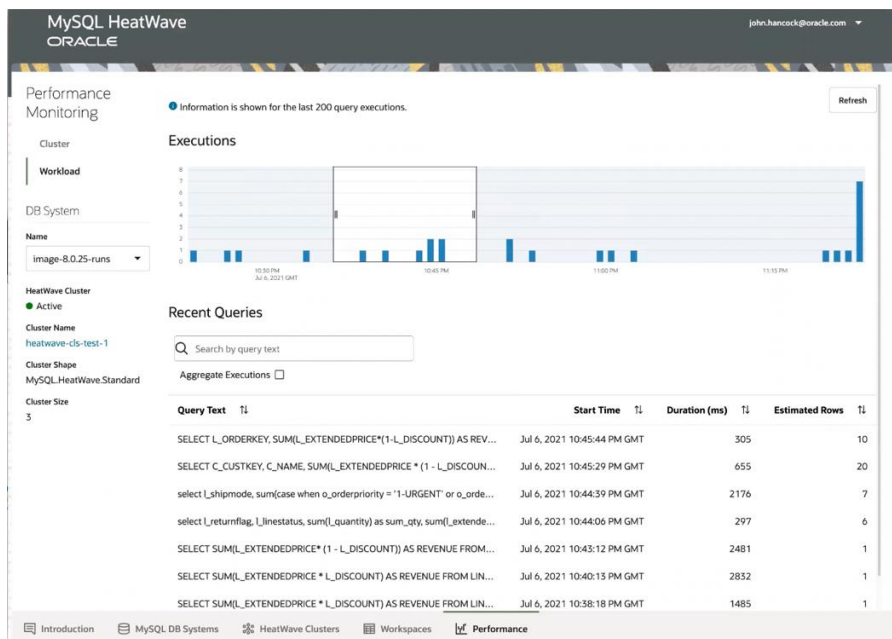
```

SELECT
  l_recomflag,
  l_linestatus,
  SUM(l_quantity) AS sum_qty,
  SUM(l_extendedprice) AS sum_base_price,
  SUM(l_extendedprice * (1 - l_discount)) AS sum_disc_price,
  SUM(l_extendedprice * (1 - l_discount) * (1 + l_tax)) AS sum_charge,
  AVG(l_quantity) AS avg_qty,
  AVG(l_extendedprice) AS avg_price,
  AVG(l_discount) AS avg_disc,
  COUNT(*) AS count_order
FROM
  tpch_lineitem
WHERE
  l_shipdate <= DATE '1998-12-01' - INTERVAL '100' DAY
GROUP BY
  l_recomflag,
  l_linestatus
ORDER BY
  l_recomflag,
  l_linestatus;
  
```

Query Results
Query completed (3.170 seconds)

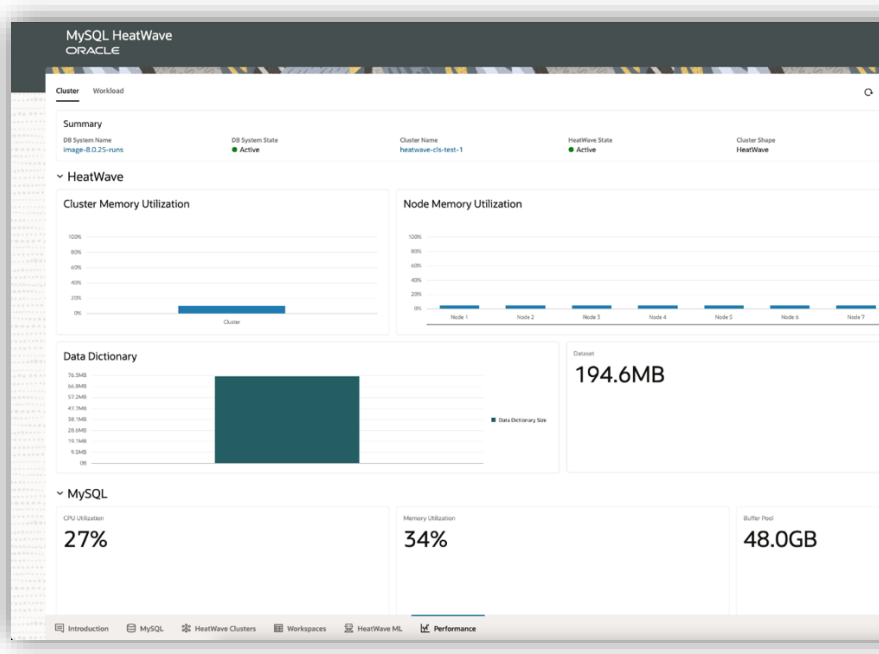
l_recomflag	l_linestatus	sum_qty	sum_base_price	sum_disc_price	sum_charge	avg_qty	avg_price	avg_disc	count_order
A	F	3865494247.00	57975897720339.54	55073056575227.9084	57280139974700.609949	25.499902	38237.095670	0.050001	1516221269
N	F	NULL	1515655546292.85	1437964915505.1633	1495505427205.797021	25.501252	38237.820214	0.050007	39585299
N	O	76132027266.00	NULL	108452120147020.6196	112790257632079.0a8476	NULL	38237.363905	0.049999	298566022
R	F	3866342979.00	57975227940301.93	55075481003326.9393	NULL	25.500078	38237.067210	0.049999	1516204881
A	F	3865494247.00	57975897720339.54	55073056575227.9084	57280139974700.609949	25.499902	38237.095670	0.050001	1516221269
N	F	NULL	1515655546292.85	1437964915505.1633	1495505427205.797021	25.501252	38237.820214	0.050007	39585299

The MySQL HeatWave console provides a query editor to ease the customer's interaction with their database system by eliminating the need to go back and forth between the console and an external MySQL client for resource and data management, respectively. Customers can write and execute queries on MySQL HeatWave and view the query results through the query editor, while monitoring the state of the database, all in the same console.



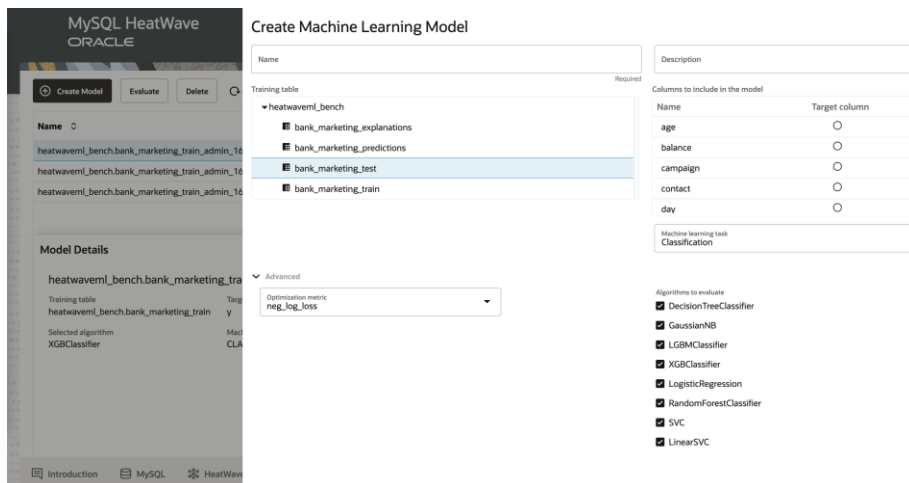
In addition, MySQL HeatWave workload monitoring summarizes the results of the most recent queries executed along with their execution time and the number of rows in the result set, allowing the customers to have access to the recent history of activities in the console.

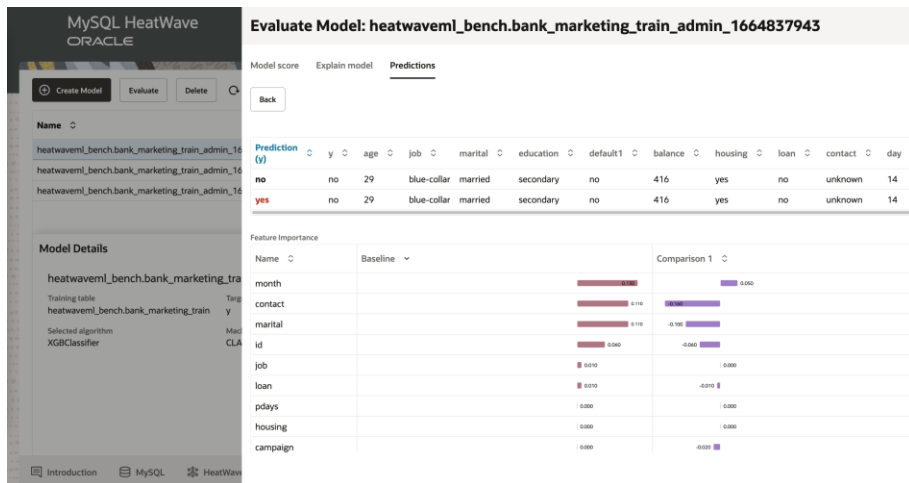
Performance Monitoring



The MySQL HeatWave console allows users to monitor the overall and per-node utilization of the hardware resources such as CPU, memory, and storage, as well as a detailed breakdown of the resource consumption like data dictionary size, buffer pool size and database connections.

HeatWave AutoML Model Creation, Prediction, and Explanation





The MySQL HeatWave console on AWS allows customers to train machine learning classification and regression models, to use the trained models for predictions, and to inspect the explanations for the models and predictions. Users can include/exclude machine learning algorithms to train, include/exclude the columns/features to train on, specify a particular scoring metric to optimize during training or evaluate post-training, and evaluate what-if scenarios by changing the input feature values for a particular row/sample to understand how they affect the model's predictions.

Conclusion

Oracle believes in giving customers a choice. By making MySQL HeatWave natively available on AWS, customers can very easily benefit from the only cloud database service that combines transactions, analytics, and machine learning services into one MySQL Database, delivering real-time, secure analytics without the complexity, latency, and cost of ETL duplication—on AWS. Customers don't have to face exorbitant data egress fees charged by AWS and higher latency when accessing a database service running in Oracle's cloud. MySQL HeatWave on AWS is optimized for AWS with a superior architecture that delivers higher performance and lower cost, as demonstrated by industry-standard benchmarks.

[Try MySQL on HeatWave on AWS for free now!](#)

Connect with us

Call +1.800.ORACLE1 or visit [oracle.com](https://www.oracle.com). Outside North America, find your local office at: [oracle.com/contact](https://www.oracle.com/contact).

 blogs.oracle.com

 facebook.com/oracle

 twitter.com/oracle

Copyright © 2022, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle, Java, and MySQL are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Benchmark queries are derived from TPC-H benchmark, but results are not comparable to published TPC-H benchmark results since they do not comply with TPC-H specification.