DevLive Level Up MySQL Summit

MySQL HeatWave Overview for Every Developer

Nipun Agarwal

Senior Vice President, MySQL Database & HeatWave Oracle

MySQL HeatWave

Single database for OLTP, analytics, machine learning and datalake



MySQL HeatWave – no change to existing MySQL applications



Price-Performance comparison with analytic databases

18x better than Redshift, 33x better than Snowflake, 42x better than BigQuery, 17x 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 Big Query is based on monthly flat rate commitment. For Azure Synapse is based on 1 year reserved pricing
- 4 Copyright © 2022 Oracle and/or its affiliates. | Oracle Confidential Includes non-disclosure information that must be covered by a current NDA agreement.

1. Designed for performance and scale

Enabled by a massively partitioned architecture



- High-fanout workload-aware partitioning
- Optimized for cache size and memory hierarchy of underlying hardware
- New algorithms help process these partitions at a high rate

2. Optimized for Commodity Cloud Infrastructure

Help provide compelling price performance



- Compute flex shapes, balancing compute with memory bandwidth, optimized for underlying architecture
- Object store bandwidth, RPS, block size, compression
- Network adaptive data flow, pipelining compute with communication

3. MySQL Autopilot: machine learning powered automation For OLTP, OLAP, Lakehouse workloads



Good Scale factor for *end-to-end* **queries** TPCH, TPCDS



Real-time elasticity

Zero downtime, system balanced after elasticity





AWS Redshift

10 –<u>15 minute</u> query blackout and manual balancing required



Stop & Resume in constant time

Autopilot related statistics are saved during stop, application resumes without any degradation



Resume Cluster



In-database Machine learning with MySQL HeatWave

Support for training, prediction, explanations



Explanation

- Fully automated pipeline
- Meta learned proxy models very performant
- Optimized to run on commodity CPUs no additional cost
- All models can be explained

ML tasks supported

- 1. Classification
- 2. Regression
- 3. Time series forecasting (univariate, multi-variate)
- 4. Unsupervised anomaly detection
- 5. Recommender systems

5 simple SQL calls: ML_TRAIN, ML_MODEL_LOAD, ML_EXPLAIN_ROW, ML_PREDICT_ROW, ML_SCORE

HeatWave AutoML is *faster* and more accurate than Redshift ML 25x faster than Redshift ML

	Aco	curacy	Training Tim	ne (minutes)	
Dataset	Redshift ML	HeatWave ML	Redshift ML	HeatWave ML (2 nodes)	Speedup
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

13 Copyright © 2022 Oracle and/or its affiliates. | Oracle Confidential – Includes non-disclosure information that must be covered by a current NDA agreement.

Customers using machine learning with MySQL HeatWave

"Using Oracle HeatWave AutoML for our recommendation system development **was like having a time machine**. **We explored and analyzed data in a flash,** with no need for dedicated servers or long data training waits with Python.."

Marcello Violini, CEO Teyuto



teyuto

"Without a doubt, Oracle has helped us sell our payment default solutions to credit institutions throughout Colombia and beyond thanks to the automated machine learning engine within MySQL HeatWave..."

> Yelitza Romero, CEO Artificial Intelligence Collection System S.A.S



"We found HeatWave AutoML very innovative, easy to use, very fast and most important it is secure since the data or the model don't leave the database..."

Arvind Rajan, CEO Astute Business Solutions

Significant growth of data outside databases

- Databases are systems of record
- Data lake is the repository for other types of data
- 79 Zetabytes of data generated in 2021, 180 ZB expected in 2025 (IDC)
- In 2019, IOTs alone generated 13.6 ZB of data
- 99.5% of collected data remains unused
- More than 80% of data is unstructured



MySQL HeatWave Lakehouse

Data in object store remains in object store



Performance of MySQL HeatWave Lakehouse

Faster to load, faster to query and less expensive than other databases

400TB TPC-H (CSV format)	MySQL HeatWave Lakehouse	Redshift	Snowflake
Annual cost	\$1,589,036	\$2,261,760	\$2,242,560
Load Time	4.1 hrs	32.3 hrs	11.2 hrs
Query Time	42 sec	268 sec	712 sec

MySQL HeatWave available on OCI, AWS, Azure



estuda.com

Estuda.com achieves real-time insights

"We have 3 million students a month on our platform, and every user who accesses estuda.com has to see their assessment report in real time. MySQL HeatWave on Oracle Cloud Infrastructure delivers the high query performance and elasticity that we require, at predictable low costs."

Vitor Pereira de Freitas

Cofounder and CTO, Estuda.com

Business Challenge:

Estuda provides a platform that aims to prepare students in Brazil for higher education institutions. Its users expect fast speeds, but the company's rapid growth hampered its ability to deliver. The critical challenge for Estuda was to find a cloud-based database that was fast, cost-effective, and effortlessly and quickly scalable.

Results:

- ✓ Migrated from Google BigQuery
- ✓ Accelerated complex queries by 300X compared to BigQuery
- ✓ Reduced costs by 85% and eliminated ETL processes
- $\checkmark\,$ Onboard more students and teachers at predictable low costs

Products Used:

MySQL HeatWave



Tetris.co simplifies real-time analytics

"MySQL HeatWave dramatically reduced our AWS Aurora and Redshift cost by more than 50%. We are no longer moving data around so now we have blazing fast, real-time insights with no effort. More importantly, scalability has made our expansion plan possible, allowing us to onboard more data and new clients without impact to costs. It's a dream come true."

Pablo Lemos

Cofounder and CTO, Tetris.co

Copyright © 2023, Oracle and/or its affiliates

Business Challenge:

Tetris.co offers analytical software that unifies massive amounts of data from several sources to help Brazilian clients understand how digital marketing investments perform. It needed a data platform that is easy to use, low cost, and scalable.

Results:

- Migrated to MySQL HeatWave from AWS Aurora and Redshift and GCP BigQuery in one month
- ✓ Accelerated complex queries from minutes to milliseconds
- $\checkmark\,$ Reduced costs by more than 50% and eliminated ETL processes
- Onboard new clients of any size at no cost impact

Products Used:

MySQL HeatWave

MySQL HeatWave Summary

- 1. Fully managed MySQL database service for OLTP, analytics and machine learning
- 2. Scale out processing of data in object store
- 3. Offers best performance and price performance for query processing and machine learning
- 4. The only MySQL service with machine learning based automation (MySQL Autopilot)
- 5. Available in all commercial regions of OCI
- 6. Available in OCI, AWS, Azure

Thank You.

