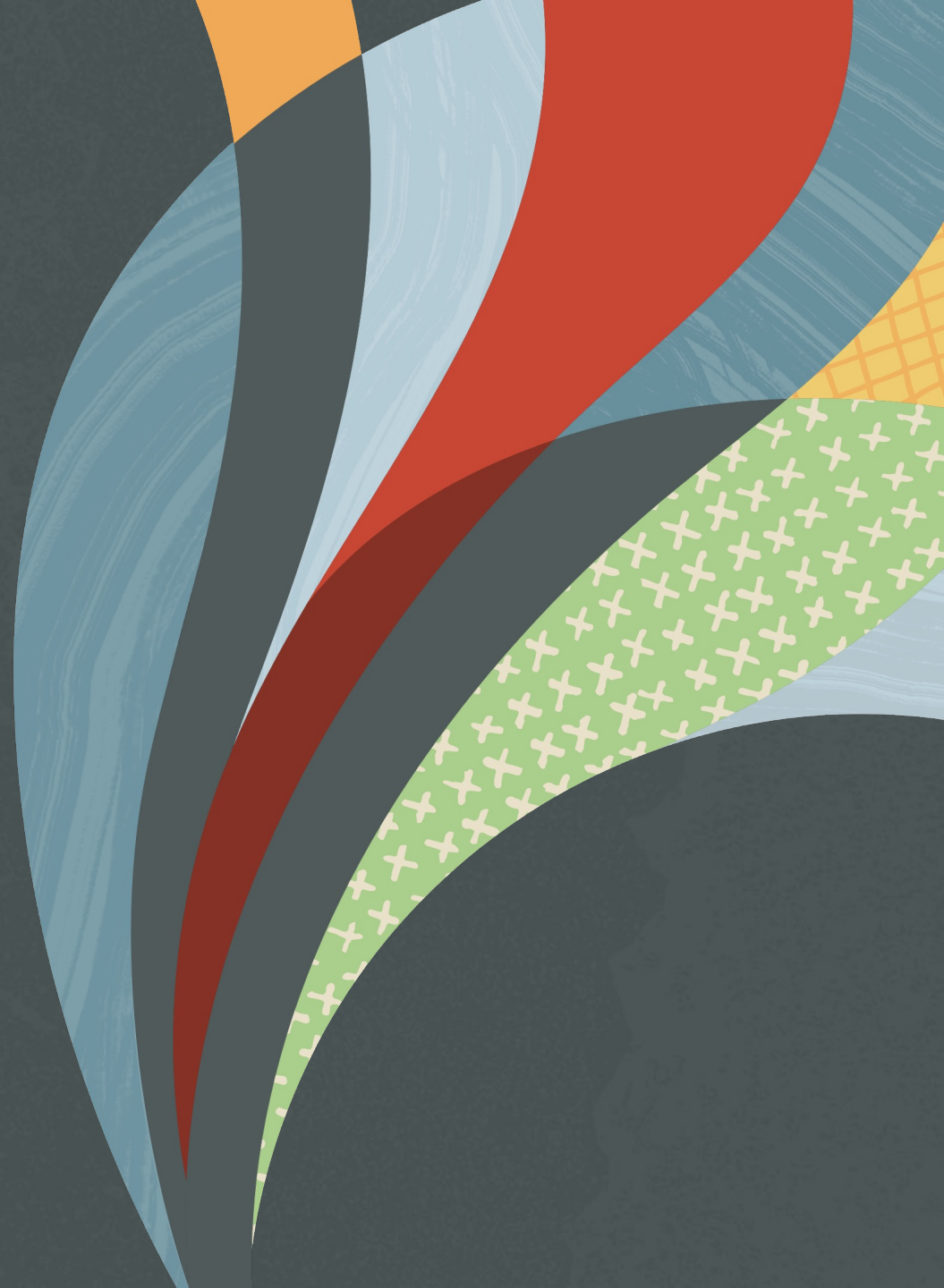


Replace Five Services with One – Using MySQL HeatWave Natively on AWS

Mandy Pang

Senior Principal Product Manager

MySQL HeatWave



Agenda

- Complexity of using AWS services
- MySQL HeatWave on AWS
- Customer success stories

Data Platform Complexity

- **You start** with an OLTP database application

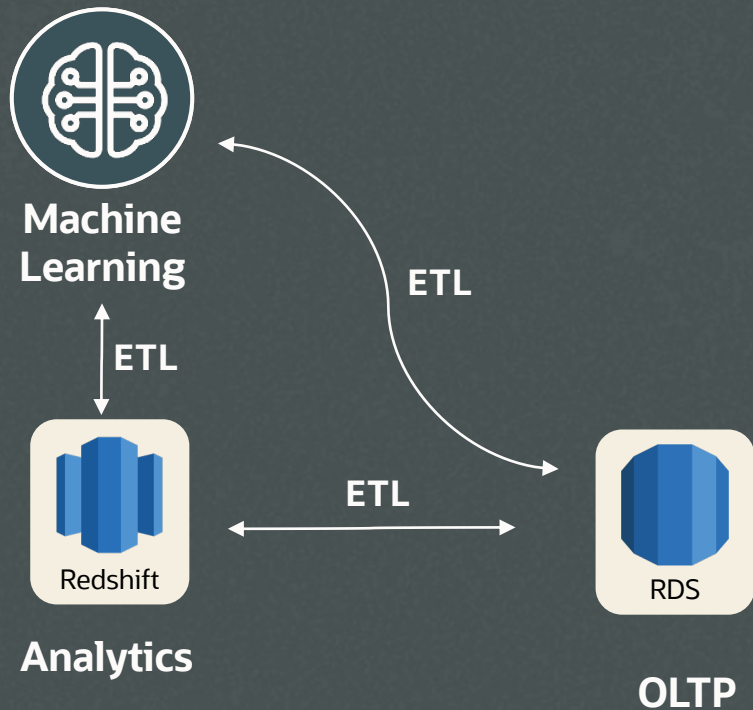


Data Platform Complexity

- **You start** with an OLTP database application
- **Analytics** will give LOB managers valuable insights

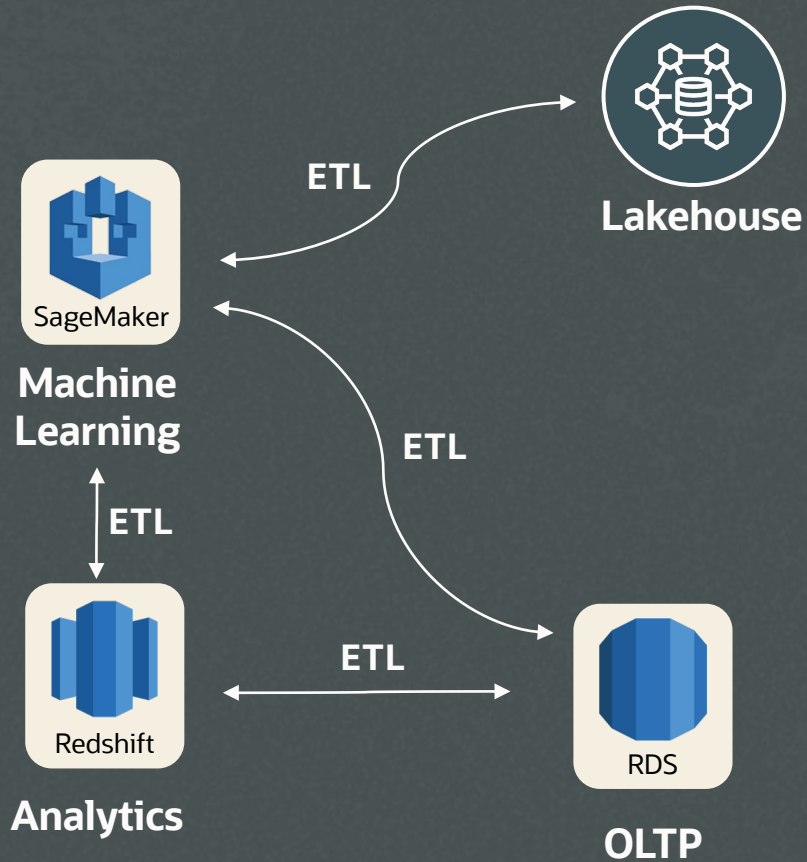


Data Platform Complexity



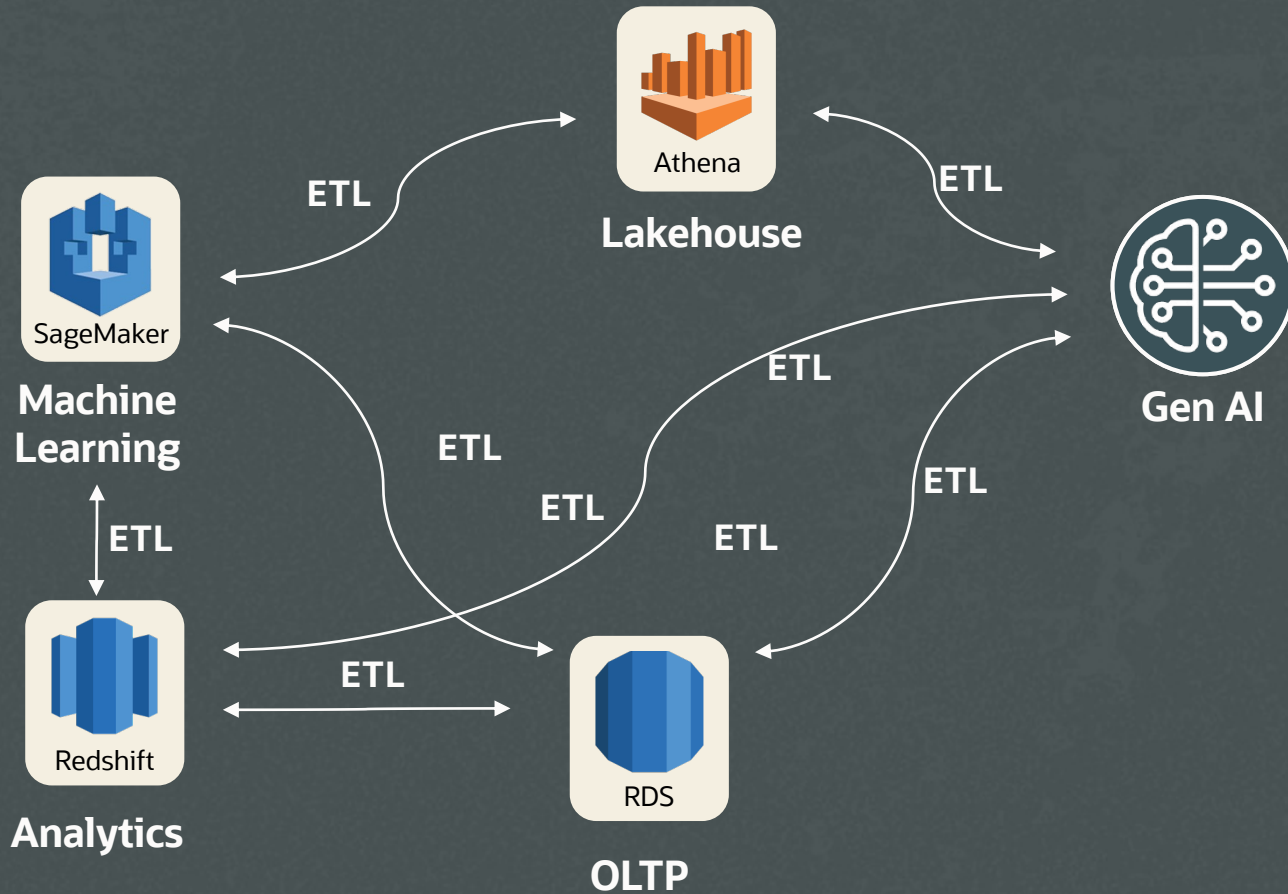
- **You start** with an OLTP database application
- **Analytics** will give LOB managers valuable insights
- **Machine learning** predictions will improve the customer experience

Data Platform Complexity



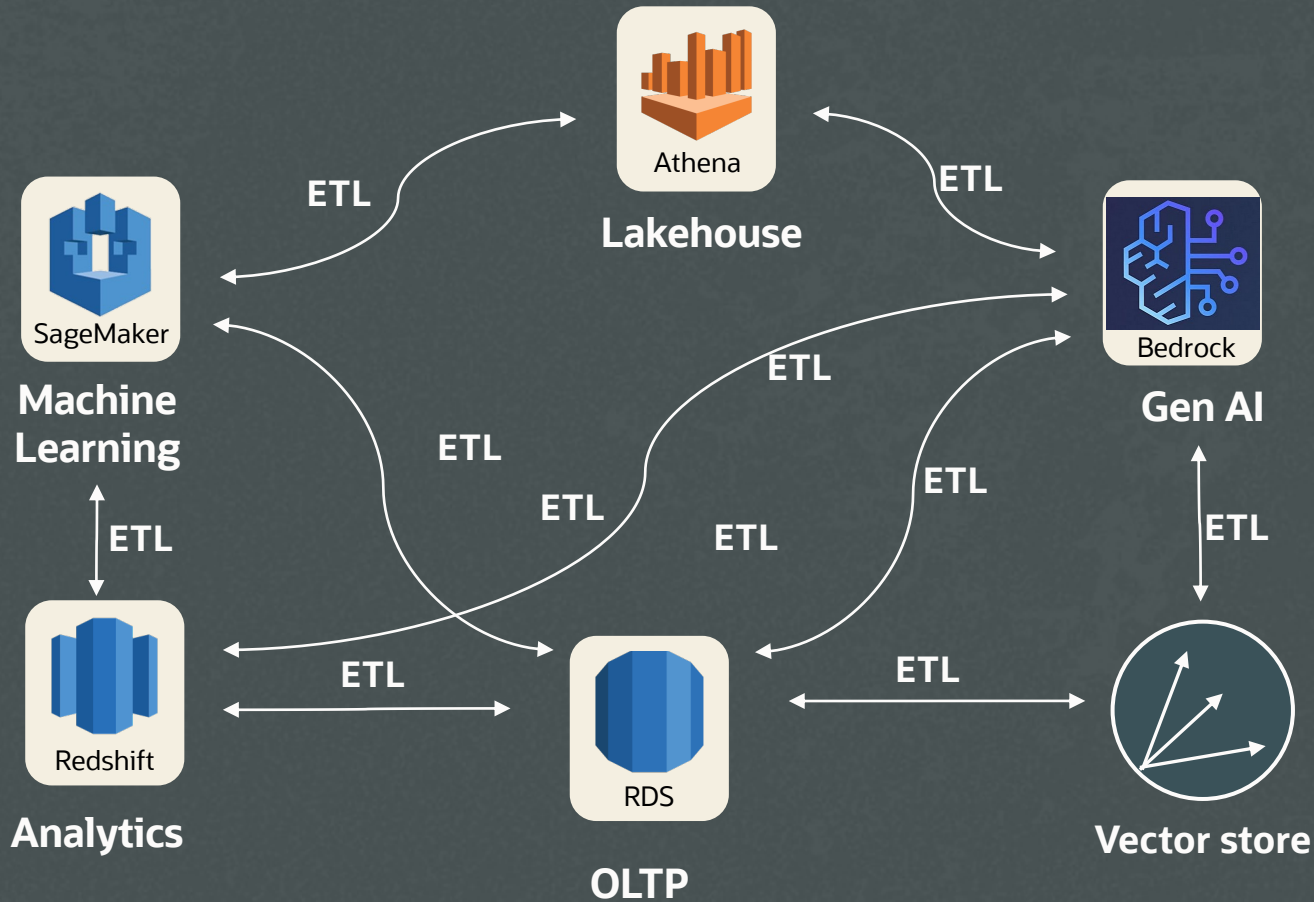
- **You start** with an OLTP database application
- **Analytics** will give LOB managers valuable insights
- **Machine learning** predictions will improve the customer experience
- **Lakehouse** will deliver insights into unstructured data

Data Platform Complexity



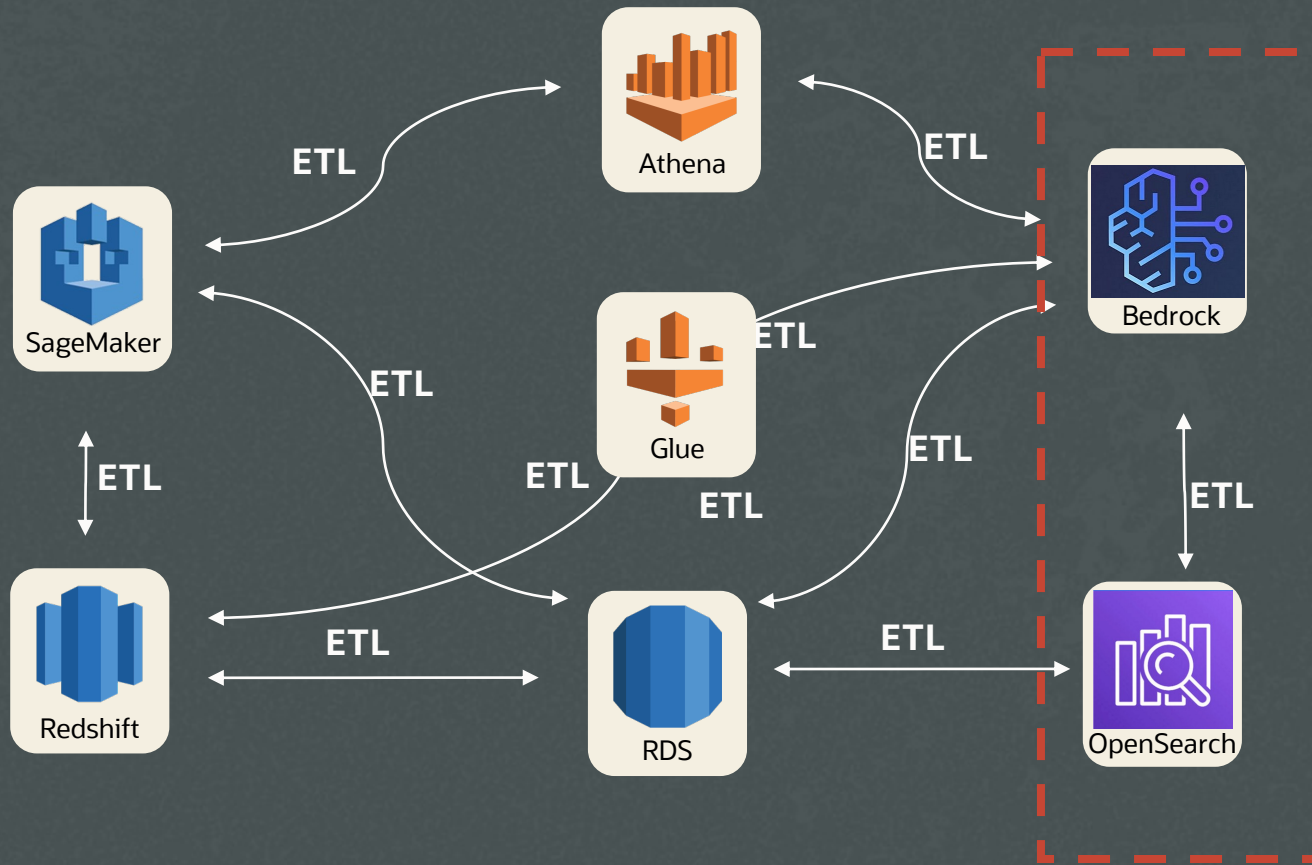
- **You start** with an OLTP database application
- **Analytics** will give LOB managers valuable insights
- **Machine learning** predictions will improve the customer experience
- **Lakehouse** will deliver insights into unstructured data
- **Generative AI** will give results in natural language

Data Platform Complexity



- **You start** with an OLTP database application
- **Analytics** will give LOB managers valuable insights
- **Machine learning** predictions will improve the customer experience
- **Lakehouse** will deliver insights into unstructured data
- **Generative AI** will give results in natural language
- **Vector store** provides context to LLM for more relevant results

Data Platform Complexity at AWS

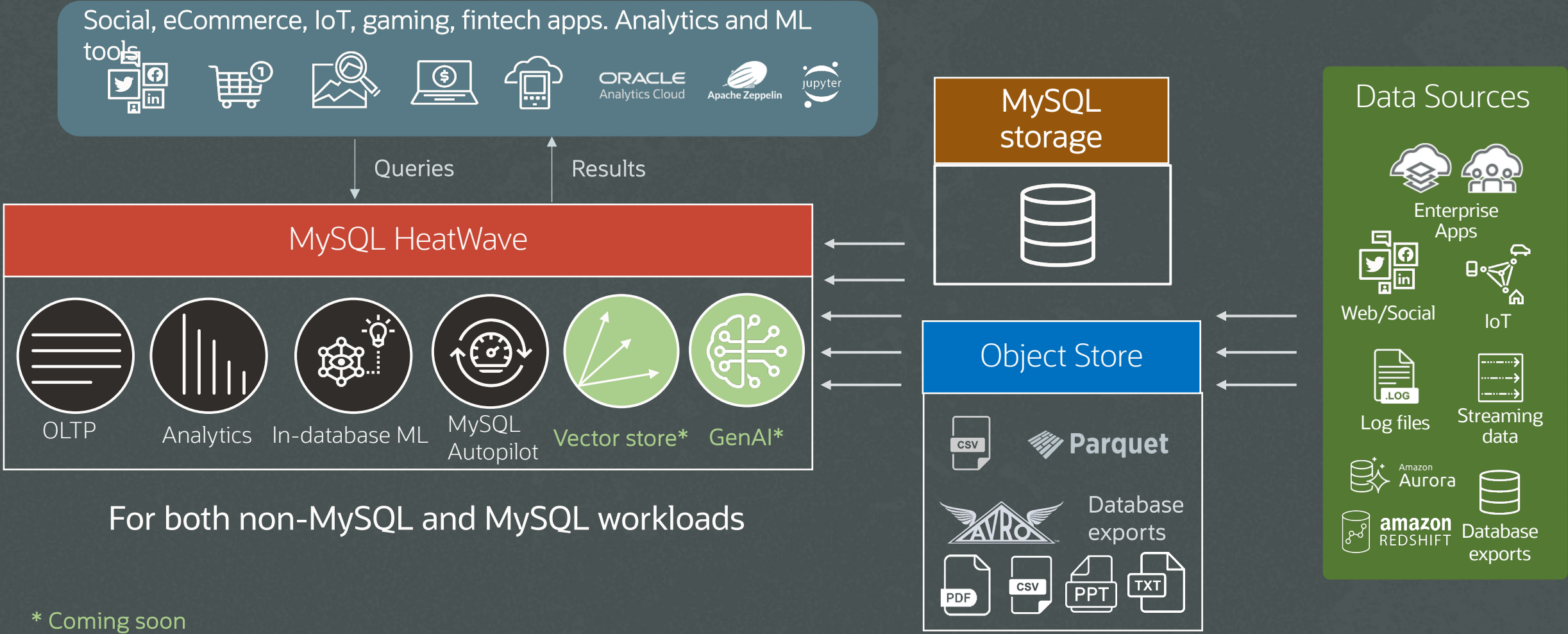


5 Separate Cloud Services

- Complex ETL processes
- Stale and obsolete data
- Difficult to maintain
- Security vulnerabilities
- Requires specialized skills

MySQL HeatWave overview

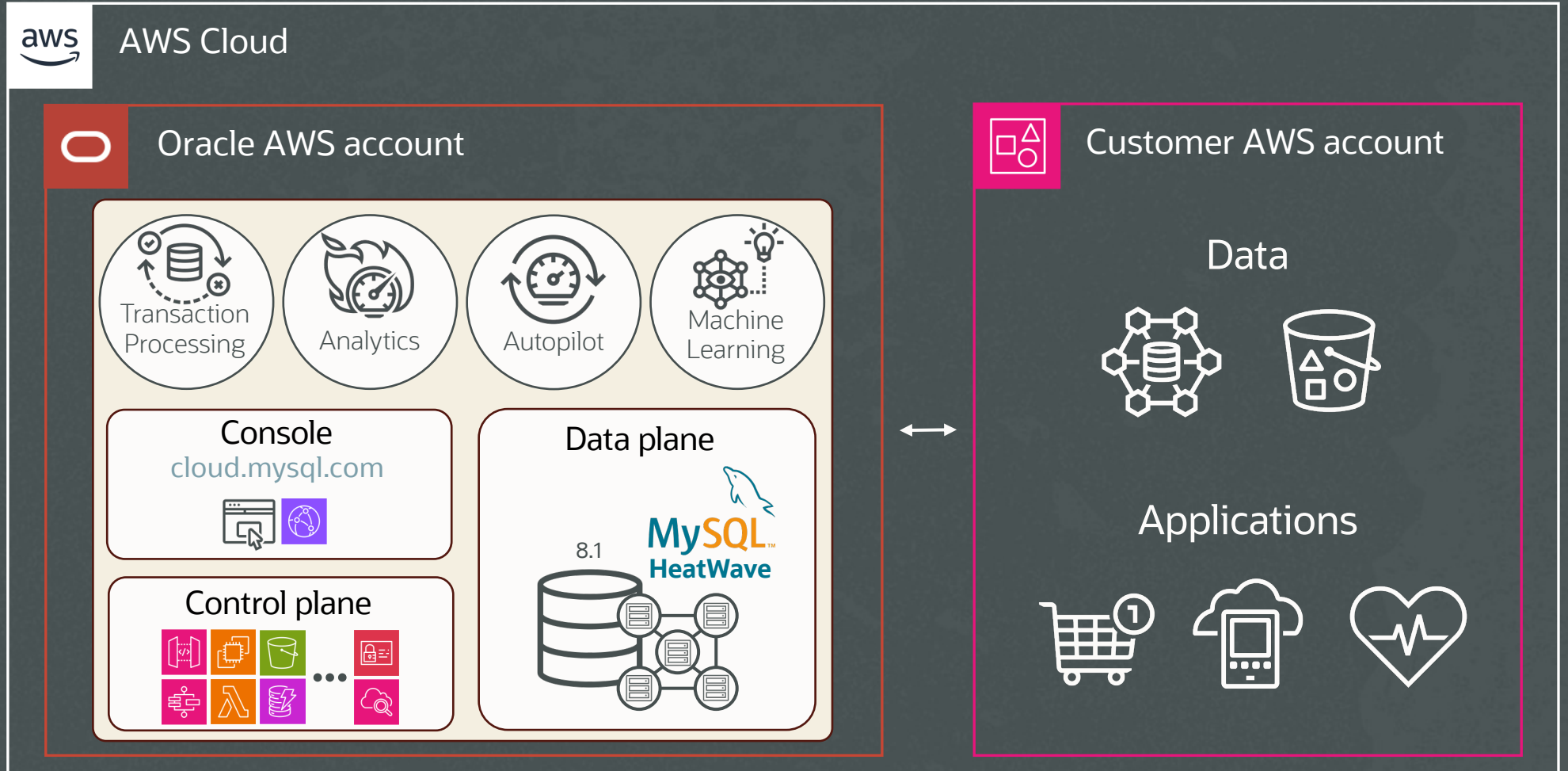
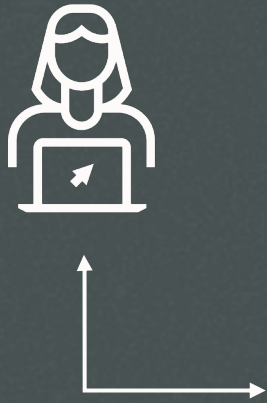
Transactions, real-time analytics, machine learning and GenAI across data warehouse and data lake in one service



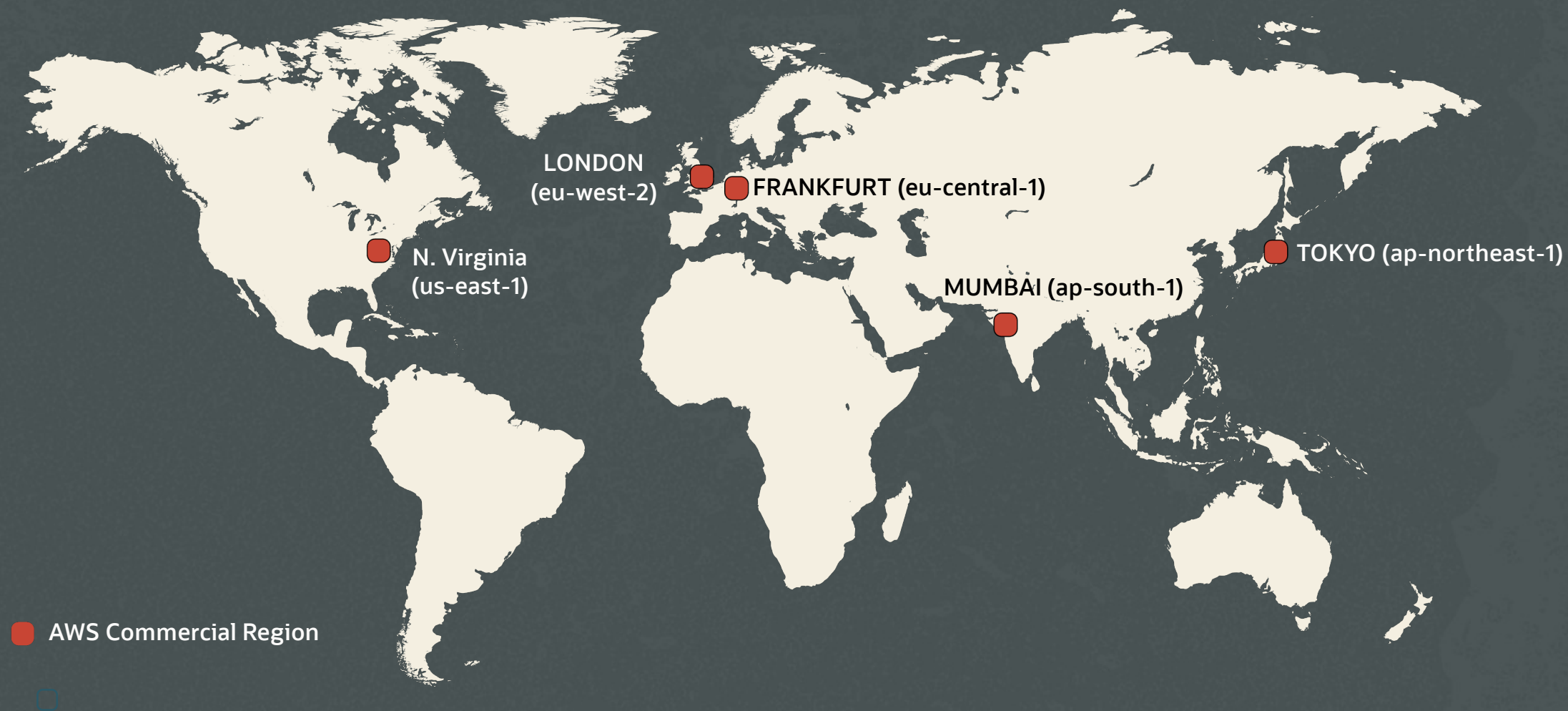
* Coming soon

MySQL HeatWave on AWS

CONSOLE, CONTROL PLANE, AND DATA PLANE RUN NATIVELY IN AWS

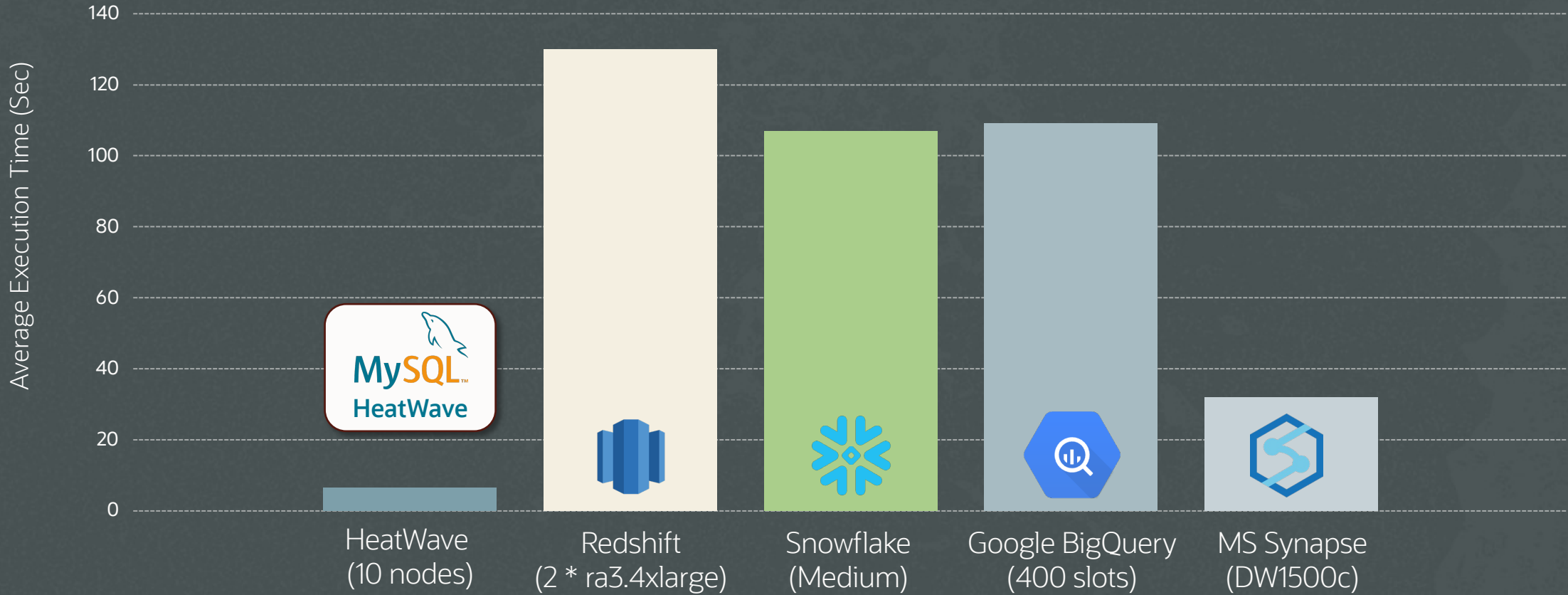


MySQL HeatWave on AWS current availability



Faster than Redshift, Snowflake, Big Query, Synapse

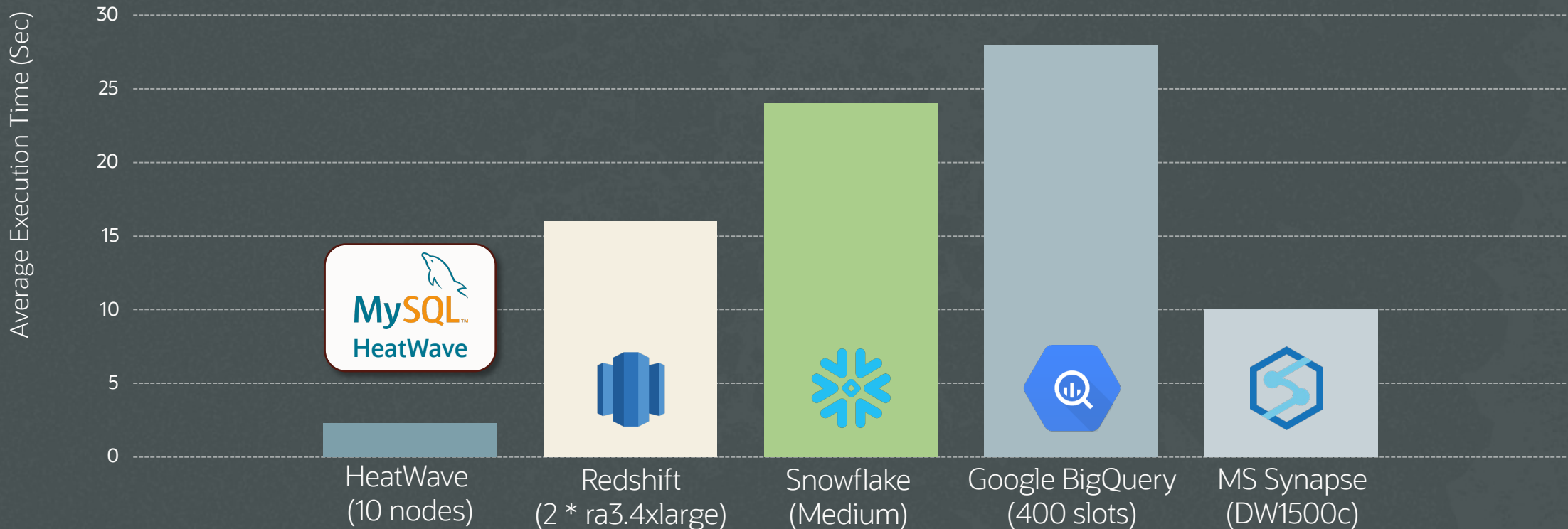
20X FASTER THAN REDSHIFT, 16X FASTER THAN SNOWFLAKE, 16X FASTER THAN BIG QUERY, 5X FASTER THAN SYNAPSE – TPC-H 4TB



* 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.
* Results from Sept 2022

Better price performance than Redshift, Snowflake, Big Query, Synapse

7X BETTER THAN REDSHIFT, 10X BETTER THAN SNOWFLAKE, 12X BETTER THAN BIG QUERY, 4X BETTER THAN SYNAPSE – TPC-H 4TB



- Only compute costs are considered above
- Pricing for Redshift is based on 1-year reserved instance, paid upfront. Snowflake is based on **standard edition**
- Pricing for Google Big Query is based on monthly flat rate commitment. Azure Synapse is based on 1-year reserved pricing

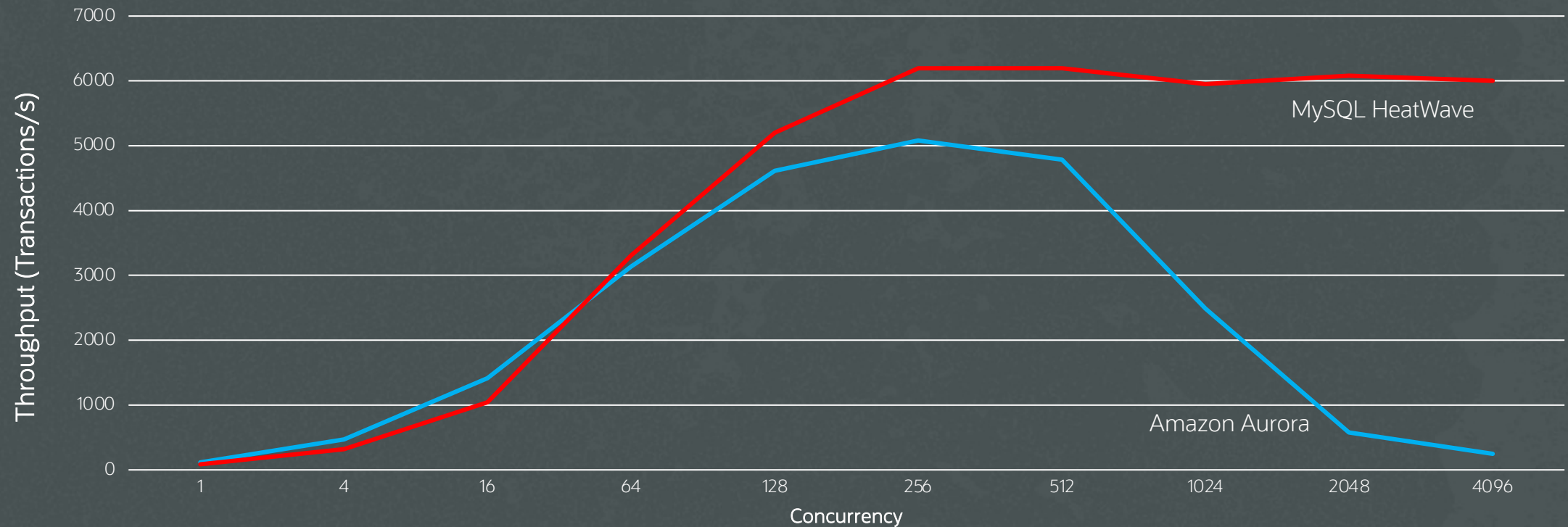
* Benchmark queries are derived from the TPC-H benchmarks, but results are not comparable to published TPC-H benchmark results since these do not comply with the TPC-H specifications

* Results from Sept 2022

MySQL HeatWave offers up to 10x better throughput than Aurora for OLTP

Auto thread pooling

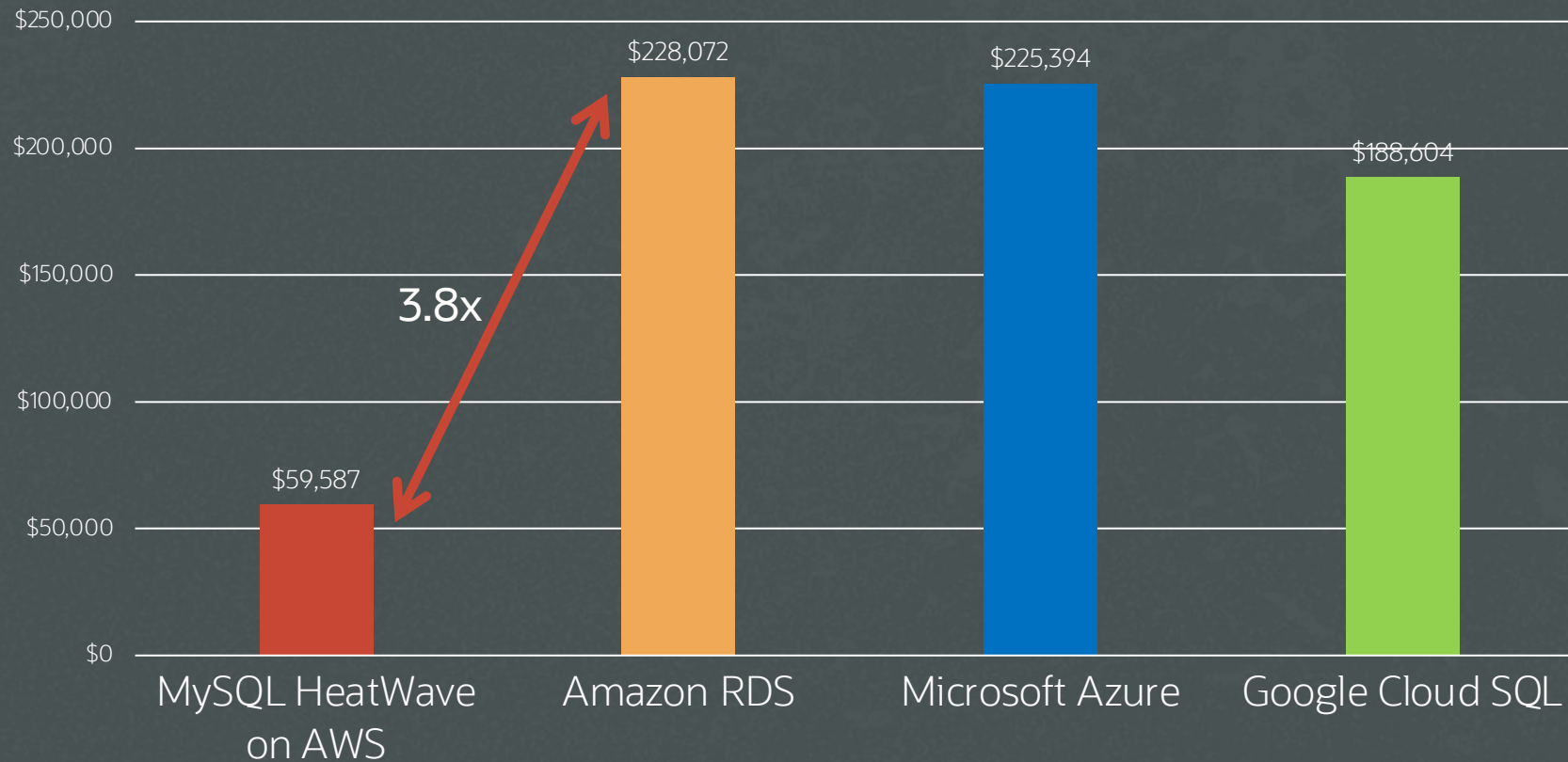
TPC-C_100W (10G, Data fits in Buffer Pool)



*Benchmark queries are derived from the TPC-C benchmarks, but results are not comparable to published TPC-C benchmark results since these do not comply with the TPC-C specifications.

Unbeatable price/performance for OLTP

MySQL HeatWave on AWS for OLTP, 1 Year TCO
200 vCPUs, 10TB Storage



MySQL HeatWave: 1 ECPU, all regions have the same price.

Amazon RDS: Intel R5 16GB/Core, AWS US East.

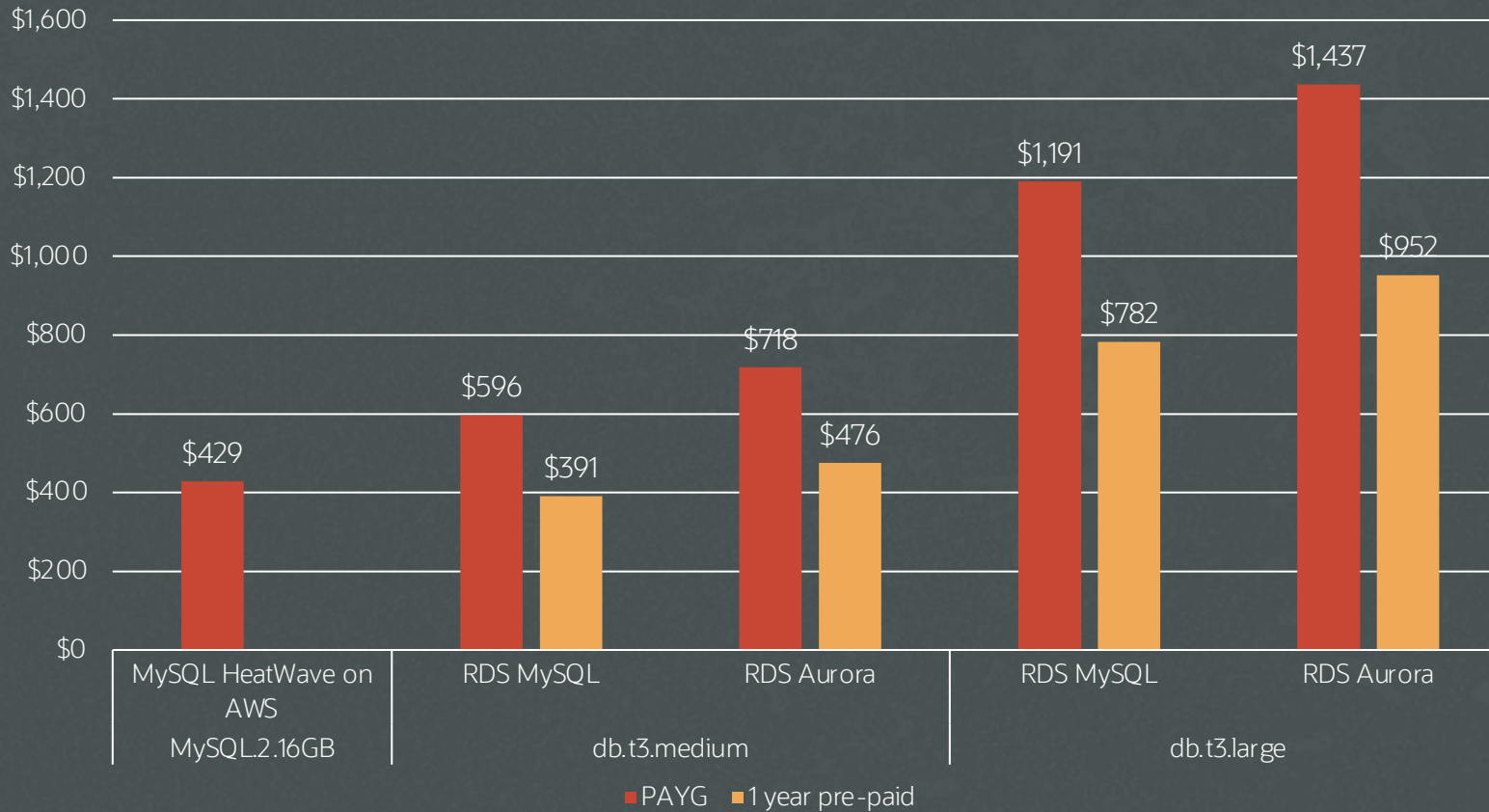
Azure: Memory Optimized Intel 20GB/Core, MS Azure US-East.

Google: High Memory N1 Standard Intel 13GB/Core, GCP Northern Virginia.

Unbeatable price/performance for small OLTP

AWS RDS/Aurora shape instances

MySQL HeatWave vs AWS RDS t3 db instance

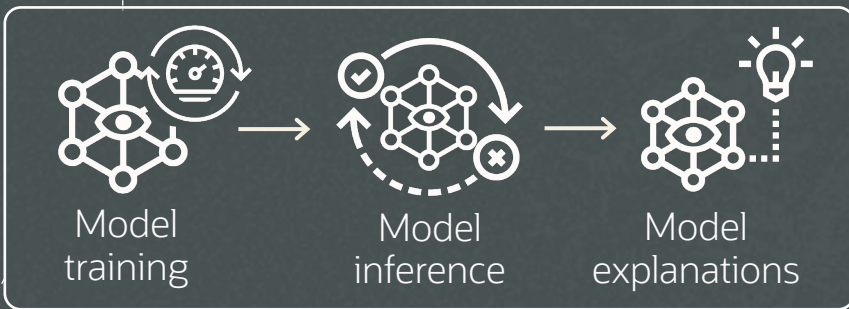
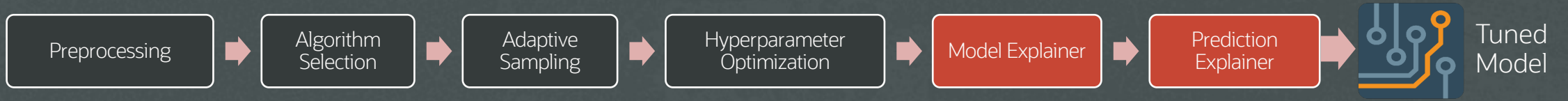


- 1.39x cheaper than RDS t3.medium
- 1.67x cheaper than Aurora t3.medium
- 2.78x cheaper than RDS t3.large
- 2.35x cheaper than Aurora t3.large

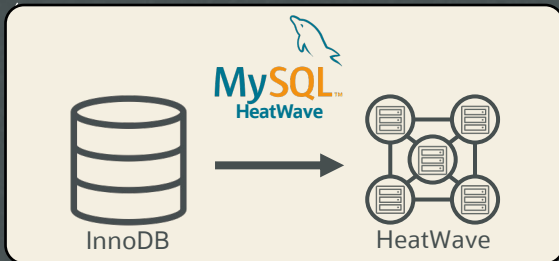
Note:

MySQL.2.16GB (2 vCPU, 16GB MEM)
RDS t3.medium (2 vCPU, 4GB MEM)
RDS t3.large (2 vCPU, 8GB MEM)

HeatWave AutoML: In-database machine learning



In-database ML



- Eliminates tedious and laborious steps
- Simple to use interface for beginner or advanced ML users
- Automatically selects algorithm and tunes it
- Explainable model behavior and predictions
- Fast training allows to quickly iterate to achieve desired outcome

Industries and use cases with HeatWave AutoML

Digital Marketing

Cost per acquisition

Targeted campaigns

Customer classification

E-Commerce

Videos for users

Lottery suggestions

Product upsell

Education

Predict student success

Monitor student behavior

HIPPA Compliance

Services

Erroneous ledger entries

Predict future losses

Predict price elasticity

FinTech

Loan default prediction

Identify loan extensions

Loan approval

Gaming

Player churn detection

Adjust game difficulty

Identify game hackers

Internet Of Things

Airport ticketing

Rain water level

Air pollution

Manufacturing

Reduce warranty claims

Defective part identification

Detect anomalies in supplies

Machine learning with HeatWave is fast, cost effective, accurate and scalable

25x

faster than
Redshift

1%

of the cost of
Redshift

MySQL HeatWave on AWS Console

Monitoring

MySQL HeatWave ORACLE

Cluster: **Workload**

Summary
 DB System Name: image-8.0.25-runs
 DB System State: **Active**
 Cluster Name: heatwave-cl-test-1
 HeatWave State: **Active**
 Cluster Shape: **HeatWave**
 Cluster Size: **3**

HeatWave
 Cluster Memory Utilization
 Data Dictionary
 MySQL
 CPU Utilization: **27%**

HeatWave Cluster
 Node Count: **3**
 Cluster Memory Snapshot

Interactive Query UI & Data Management

MySQL HeatWave ORACLE

Workspaces
 DB System: image-8.0.25-runs
 Query 1
 Run Clear HeatWave Active DB System image-8.0.25-runs user admin

```

1 SELECT
2   l_returnflag,
3   l_linestatus,
4   SUM(l_quantity) AS
5   SUM_l_extendedprice
6   SUM(l_extendedprice
7   SUM_l_extendedprice
8   AVG(l_quantity) AS
9   AVG_l_extendedprice
10  AVG(l_discount) AS
11  COUNT(*) AS count_o
12 FROM
13  tpch.LINEITEM
14 WHERE
15   l_shipdate <= DATE
16 GROUP BY
17   l_returnflag,
18   l_linestatus
19 ORDER BY
20   l_returnflag,
21   l_linestatus;
  
```

Query Results
 Query completed (5.1700 seconds)
 Results: JSON Job Info

Model Details
 Name: DEMO_CORPUS.bank_marketing_train_admin_1663723896
 State: Active
 Prediction column: y
 Training table: DEMO_CORPUS.bank_marketing_train
 Target column: y
 Selected algorithm: XGBoost
 Machine learning task: CLASSIFICATION
 Description: TBD
 Training score: None: 0

HeatWave ML

Evaluate Model: DEMO_CORPUS.
 Model score Explain model Predictions

Estimate Cluster Size with MySQL Autopilot

Refresh Estimate Estimate last refreshed on Mon, 18 Apr 2022 23:23:16 GMT

Schemas	Schema name	Tables from selected schemas	Table name
<input type="checkbox"/>	sampleDB		
<input type="checkbox"/>	sampleDB2		
<input checked="" type="checkbox"/>	tpcds_1024		
<input checked="" type="checkbox"/>	tpch1024		
<input type="checkbox"/>	tpch4096		

Name	HeatWave Cluster Memory Usage (GiB)	Tables Selected	Warnings
tpcds_1024	820.636	25 of 25	1
tpch1024	1184.211	8 of 8	0
tpch4096	0.000	0 of 8	0

Name	Warnings	Memory Size Estimate (GiB)	Rows Estimate
tpcds_1024.CALL_CENTER	0.003	6	
tpcds_1024.CATALOG_PAGE	0.003	11,471	
tpcds_1024.CATALOG_RETURNS	19.578	159,947,748	
tpcds_1024.CATALOG_SALES	304.989	1,488,072,176	
tpcds_1024.CUSTOMER	0.037	311,767	
tpcds_1024.CUSTOMER_ADDR...	0.016	156,457	
tpcds_1024.CUSTOMER_DEMO...	0.084	1,911,910	
tpcds_1024.DATE_DIM	0.009	72,540	
tpcds_1024.DBGEN_VERSION	1		
tpcds_1024.HOUSEHOLD_DEM...	0.003	7,311	
tpcds_1024.INCOME_BAND	0.003	20	
tpcds_1024.INVENTORY	0.253	12,174,979	
tpcds_1024.ITEM	0.005	17,422	
tpcds_1024.PROMOTION	0.003	303	
tpcds_1024.REASON	0.003	35	
tpcds_1024.SHIP_MODE	0.003	20	
tpcds_1024.STORE	0.003	16	

Summary

- 2004.847 (GiB) Memory required by the schemas/tables selected
- 256 (GiB) Memory provided per node
- 8 HeatWave Cluster nodes required
- 2048 (GiB) Memory provided by 8 node cluster

Apply Cluster Size Estimate Cancel

Quick start with a few clicks

Starter DB System

- Create MySQL HeatWave instance with predefined configurations and sample data pre-loaded to HeatWave

Sample Data

- Import sample schema such as airportdb and TPCH into the DB System
- Run sample analytics queries
- Evaluate MySQL HeatWave performance and features

MySQL HeatWave ORACLE

Query Editor Manage Database

Import Data Cancel

Name

Import data

Data stored in a vault

Import data into DB System

Display name
dataImport20240405113619195

Description

Source

Bring your own data Import sample data

airportdb

The airportdb database provides a small dataset that contains airline, airport, and flight status information that you can use to run sample analytic queries in MySQL HeatWave. [Learn more](#)

[Sample airportdb queries](#)

Destination schema	Size
airportdb	1.9 GiB

TPCH

The TPCH database provides a suite of business oriented data that you can use to evaluate MySQL HeatWave performance and features. [Learn more](#)

We suggest picking a size that represents your expected workload.

[Sample TPCH queries](#)

Destination schema	Size
tpch_10	10.6 GiB

Cancel Import

HeatWave AutoML console

Democratizing machine learning

- Visual interface
- Enables business analysts to build, train, run, and explain ML models. No SQL and no coding required
- Ability to explore “What-if” scenarios
- All processing is done inside the database
- No additional cost to users

The screenshot displays the MySQL HeatWave AutoML console interface. At the top, it shows the MySQL HeatWave ORACLE logo and a user profile for 'heatwave.demo@gmail.com'. Below the header, there are buttons for 'Create Model', 'Evaluate', and 'Delete'. A table lists several trained models with columns for Name, State, Prediction column, Training table, Selected algorithm, Training score, and Created. The selected model is 'model20230315193808278', which is Active and uses an XGBClassifier on the 'Bankmarketing.bank_marketing_train' table. Below the table, a 'What if' section shows a comparison between a Baseline and Comparison 1 across various features. A feature importance chart compares the Baseline and Comparison 1 models, showing the impact of each feature on the prediction. The chart shows that 'duration' has the largest impact on predicting 'no' in the comparison, while 'month' has the largest impact on predicting 'no' in the baseline. A 'Notes' section provides a summary of these findings.

Name	State	Prediction column	Training table	Selected algorithm	Training score	Created
model20230315193808278	Active	y	Bankmarketing.bank_marketing_train	XGBClassifier	neg_log_loss: -0.2034	Wed, 15 Mar 2023 19:40:29 GMT
model20230315183708908	Active	revenue	Census.census_train	GaussianNB	neg_log_loss: -0.4264	Wed, 15 Mar 2023 18:39:06 GMT
model20230315175931984	Active	revenue	Census.census_train	LGBMClassifier	neg_log_loss: -0.2891	Wed, 15 Mar 2023 18:08:13 GMT
model20230314001756513	Active	survived	Titanic.titanic_train	XGBClassifier	neg_log_loss: -0.4418	Tue, 14 Mar 2023 00:20:59 GMT
model202303131001507826	Active	revenue	Census.census_train	GaussianNB	neg_log_loss: -0.4264	Tue, 31 Jan 2023 00:17:20 GMT

Dataset	Prediction (y)	age	job	marital	education	default1	balance	housing	loan	contact	day	month	duration	campaign	pday	
Baseline	no	yes	40	services	single	secondary	no	224	yes	no	cellular	17	apr	875	2	-1
Comparison 1	no	yes	40	services	single	secondary	no	224	yes	no	cellular	17	apr	89	2	-1

Name	Baseline	Comparison 1
month	0.110	0.0100
day	0.070	0.0100
housing	0.040	0.0300
age	0.030	0.0100
poutcome	0.010	0.0000
pdays	0.000	0.0000
contact	0.000	0.0100
duration	-0.440	-0.0200
previous	-0.040	0.0100
balance	-0.030	-0.0100

Notes:
Baseline:
month (apr) had the largest impact towards predicting no, whereas duration (875) contributed the most against predicting no.
Comparison:
housing (yes) had the largest impact towards predicting no, whereas duration (89) contributed the most against predicting no.



Aiwifi migrates from Amazon RDS to MySQL HeatWave on AWS

“With MySQL HeatWave’s incredible performance and built-in machine learning, we knocked down previous barriers to growth. Aiwifi estimates that MySQL HeatWave replaced up to 5 external systems. Making HeatWave available on multiple cloud platforms is a very smart move by Oracle.”

Eric Aguilar, CEO & CTO, Aiwifi

Challenge:

Aiwifi is a Mexican company developing Wi-Fi solutions that connect shoppers to websites through customized captive portals. Its value proposition is to gather valuable customer data by tracking user profiles and activity. Upon starting up in 2019, Aiwifi chose AWS as its platform and Amazon RDS as the backend database. However, as the business rapidly grew and generated heavy data loads, the lack of performance became a bottleneck for sustained growth, and database costs became a heavy challenge. In 2023, it migrated from Amazon RDS to MySQL HeatWave running natively inside AWS.

Results:

- ✓ Queries ran 13X faster and loading time on captive portals dropped by 50%, allowing Aiwifi to quickly onboard new customers without added costs
- ✓ Costs were reduced by 50%. The MySQL HeatWave high performance allowed using a smaller instance and high data egress fees were eliminated
- ✓ MySQL HeatWave efficiently handles complex queries on more than 40 million records to provide real-time analytics dashboards
- ✓ The need for query optimization was eliminated, allowing Aiwifi’s developers to focus on building machine learning models with HeatWave AutoML
- ✓ ML is used to segment their user base and create more personalized marketing content as well as to predict offers that could be of interest to different customer segments

Products Used:

[Oracle MySQL HeatWave](#)

Read [story](#)



Software company utilizing ML/GenAI

Company: Provides Automation of ITSM & GRC in an Integrated Platform

Use Cases:

1. Pre-configured processes and workflows eliminating spreadsheets and manual work
2. Maximum visibility and data insights allows users to correlate, analyze, and remediate issues
3. Flexible Platform that can scale and simplify existing stack

Solution: Leveraging HeatWave to Automate IT & Security Management

50% Reduction in ML Activity

Reduction in Data Cleaning, Model Selection, Model Tuning and Training Time

15-25% Performance Improvement

Using Auto-Indexing, Auto-ML and Javascript
Move processing closer to Data to improve latency (Javascript)

Tech. Consolidation

Seamlessly extending relational model to support OLAP/ML/AI workload
Reduction in maintaining and deploying multiple Technology stack and training



Dubai-based online aggregator connecting thousands of its users to their favorite restaurants, making online ordering easier, reliable and convenient.

Use Case

- Predict food deliver time.
- Suggest food / restaurant based on past actions.
- Summarize menu for the selected restaurant.

Model Type

- Regression, Recommendation, Generative AI

Results

- Developed ML models in days that would have otherwise taken months
- Database developers were able to build the models without ML expertise
- Simplified infrastructure with no complex ETL to manage and one platform providing OLTP, Analytics, ML and GenAI
- Consistent interface across various ML model types simplified learning for Eat Easy development team

MySQL HeatWave on AWS

Combine FIVE AWS services into ONE

- MySQL HeatWave runs natively on AWS, optimized for AWS infrastructure
- Data doesn't leave AWS – no data egress costs, and avoids compliance approvals
- Lowest latency access to MySQL HeatWave
- Tight integration with the AWS ecosystem – S3, CloudWatch, PrivateLink
- Easier migration from other databases (e.g., Amazon Aurora, Redshift, Snowflake)