Heatwave Release Notes

Abstract

This document contains release notes for the changes in each release of HeatWave.

For additional HeatWave documentation, see MySQL HeatWave User Guide.

Updates to these notes occur as new product features are added, so that everybody can follow the development process. If a recent version is listed here that you cannot find on the download page (https://dev.mysql.com/downloads/), the version has not yet been released.

The documentation included in source and binary distributions may not be fully up to date with respect to release note entries because integration of the documentation occurs at release build time. For the most up-to-date release notes, please refer to the online documentation instead.

For legal information, see the Legal Notices.

For help with using MySQL, please visit the MySQL Forums, where you can discuss your issues with other MySQL users.

Document generated on: 2022-05-19 (revision: 24727)

Table of Contents

Preface and Legal Notices .................................................................................................................. 1
Changes in HeatWave ........................................................................................................................ 3
Changes in HeatWave 8.0.28-u3 (2022-04-19, General Availability) ........................................... 3
Changes in HeatWave 8.0.28-u2 (2022-03-29, General Availability) ........................................... 3
Changes in HeatWave 8.0.28-u1 (2022-02-15, General Availability) ........................................... 4
Changes in HeatWave 8.0.27-u3 (2021-12-15, General Availability) ........................................... 4
Changes in HeatWave 8.0.27-u2 (2021-12-07, General Availability) ........................................... 4
Changes in HeatWave 8.0.26-u2 (2021-09-21, General Availability) ........................................... 5
Changes in HeatWave 8.0.26-u1 (2021-08-10, General Availability) ........................................... 5
Changes in HeatWave 8.0.26 (2021-07-23, General Availability) .................................................. 6
Changes in HeatWave 8.0.25 (2021-05-11, General Availability) .................................................. 7
Changes in HeatWave 8.0.24 (2021-04-20, General Availability) .................................................. 8
Changes in HeatWave 8.0.23-u2 (2021-03-15, General Availability) ........................................... 8
Changes in HeatWave 8.0.23-u1 (2021-02-09, General Availability) ........................................... 9
Index .................................................................................................................................................. 10

Preface and Legal Notices

This document contains release notes for the changes in each release of HeatWave.

Legal Notices

Copyright © 1997, 2022, Oracle and/or its affiliates.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted
in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs (including any operating system, integrated software, any programs embedded, installed or activated on delivered hardware, and modifications of such programs) and Oracle computer documentation or other Oracle data delivered to or accessed by U.S. Government end users are "commercial computer software" or "commercial computer software documentation" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, reproduction, duplication, release, display, disclosure, modification, preparation of derivative works, and/or adaptation of i) Oracle programs (including any operating system, integrated software, any programs embedded, installed or activated on delivered hardware, and modifications of such programs), ii) Oracle computer documentation and/or iii) other Oracle data, is subject to the rights and limitations specified in the license contained in the applicable contract. The terms governing the U.S. Government’s use of Oracle cloud services are defined by the applicable contract for such services. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

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

Intel and Intel Inside are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Epyc, and the AMD logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.

This documentation is NOT distributed under a GPL license. Use of this documentation is subject to the following terms:

You may create a printed copy of this documentation solely for your own personal use. Conversion to other formats is allowed as long as the actual content is not altered or edited in any way. You shall not publish or distribute this documentation in any form or on any media, except if you distribute the documentation in a manner similar to how Oracle disseminates it (that is, electronically for download on a Web site with the software) or on a CD-ROM or similar medium, provided however that the documentation is disseminated
Changes in HeatWave

Changes in HeatWave 8.0.28-u3 (2022-04-19, General Availability)

Functionality Added or Changed

• \texttt{expr IN (value,...)} comparisons, where the expression is a single value and compared values are constants of the same data type and encoding, have been optimized. For example, the following \texttt{IN()} comparison has been optimized:

\begin{verbatim}
SELECT * FROM Customers WHERE Country IN ('Germany', 'France', 'Spain');
\end{verbatim}

• Tables that have become stale due to a change propagation failure resulting from an out-of-code error are now automatically reloaded. A check for stale tables is performed periodically when the HeatWave Cluster is idle. Previously, identifying and reloading stale tables was a manual process. See Change Propagation.

Changes in HeatWave 8.0.28-u2 (2022-03-29, General Availability)

• HeatWave Machine Learning

HeatWave Machine Learning

• HeatWave customers now have access to HeatWave Machine Learning (ML), which is a fully managed, highly scalable, cost-efficient, machine learning solution for data stored in MySQL. HeatWave ML provides a simple SQL interface for training and using predictive machine learning models, which can be used by novice and experienced ML practitioners alike. With HeatWave ML, you can train a model with a single call to an SQL routine. Similarly, you can generate predictions with a single \texttt{CALL} or \texttt{SELECT} statement which can be easily integrated with your applications.

With HeatWave ML, data and models never leave the MySQL Database Service, saving you time and effort while keeping your data and models secure. HeatWave ML is optimized for HeatWave shapes and scaling, and all HeatWave ML processing is performed on the HeatWave Cluster. ML computation is distributed among HeatWave nodes, taking advantage of HeatWave's scalability and massively parallel
processing capabilities. For more information about HeatWave’s machine learning capabilities, see HeatWave ML.

Functionality Added or Changed

- HeatWave now compresses data as it is loaded, which permits HeatWave nodes to store more data. More data per node reduces costs by minimizing the size of the HeatWave Cluster required to store your data. Data compression is enabled by default but can be disabled at runtime using the rapid_compression session variable. For more information, see Data Compression.

Changes in HeatWave 8.0.28-u1 (2022-02-15, General Availability)

Functionality Added or Changed

- HeatWave now supports up to 1017 columns for base relations (tables as loaded into HeatWave), and up to 1800 columns for intermediate relations (intermediate tables used during query processing). The maximum column width for base relations and intermediate relations was increased to 65532 bytes. See Column Limits.

Changes in HeatWave 8.0.27-u3 (2021-12-15, General Availability)

Functionality Added or Changed

- Support was added for the `CONVERT_TZ()` and `LAST_DAY()` functions, which are used to manipulate temporal values.

- The HeatWave Cluster recovery process was optimized to avoid applying a large volume of changelogs during recovery. Snapshots are now taken when the volume of changelogs and the time required to apply those changes exceed specific thresholds, and recovery from Object Storage is performed using those snapshots.

- HeatWave now supports automatic data reload when the HeatWave Cluster is restarted. Previously, when a HeatWave Cluster was stopped by a stop or restart action, data had to be reloaded manually after the cluster was restarted. Now, when starting or restarting a HeatWave Cluster, data that was previously loaded is reloaded automatically. Data changes that occur on the DB System while the HeatWave Cluster is offline are included in the reloaded data.

  Automatic data reload does not occur if the HeatWave Cluster was stopped as a result of a stop or restart action performed on the DB System to which the HeatWave Cluster is attached. In this case, data loaded in the HeatWave Cluster must be reloaded manually after the HeatWave Cluster is restarted.

Changes in HeatWave 8.0.27-u2 (2021-12-07, General Availability)

Functionality Added or Changed

- HeatWave now supports querying views. See Using Views.

- Bloom filter join optimizations were introduced. For HeatWave queries that join large and small relations, bloom filters reduce the amount of data processed by early filtering and the amount of memory used during query processing.

- The `rpd_query_stats` table, which stores HeatWave query history (compilation and execution statistics), now stores data for the last 1000 executed queries. Previously, data was stored for the last 200 queries.

  The following columns were added to the `performance_schema.rpd_tables` table:
• **SIZE_BYTES**: The amount of data loaded per table, in bytes.
• **QUERY_COUNT**: The number of queries that referenced the table.
• **LAST_QUERIED**: The timestamp of the last query that referenced the table.
• **LOAD_END_TIMESTAMP**: The load completion timestamp for the table.

The following column was added to the `performance_schema.rpd_columns` table:

• **DICT_SIZEBYTES**: The dictionary size per column, in bytes.

The following column was added to the `performance_schema.rpd_nodes` table:

• **BASEREL_MEMORY_USAGE**: The base relation memory footprint per node.

The `rapid_query_stats` and `rpdc_exec_stats` tables are now synchronized. If a query record is removed from the `rapid_query_stats` table, it is also removed from the `rpdc_exec_stats` table.

**Changes in HeatWave 8.0.26-u2 (2021-09-21, General Availability)**

**Functionality Added or Changed**

• The following function support was added:
  
• `YEARWEEK(date)`, `YEARWEEK(date, mode)`
  
• The mode argument for the two-argument form of the `WEEK()` function: `WEEK(date[, mode])`

• `MAKEDATE()`

• “Zero” handling for dates such as ‘2001-11-00’ was implemented for `WEEK()`, `YEARWEEK()`, and `MAKEDATE()` functions.

• `CAST()` of `FLOAT` and `DOUBLE` values to `DECIMAL`

(Bug #33163625, Bug #33138534)

• The new `hw_data_scanned` global status variable tracks the total cumulative megabytes scanned by successfully executed HeatWave queries.

The number of megabytes scanned by an individual HeatWave query can be obtained by querying the `performance_schema.rpd_query_stats` table.

An estimated number of megabytes scanned by an individual query can be obtained by running the query with `EXPLAIN` and querying the `performance_schema.rpd_query_stats` table.

For more information, see [Tracking Scanned Data](#).

**Changes in HeatWave 8.0.26-u1 (2021-08-10, General Availability)**

• **HeatWave Network Layer**

• **HeatWave Data Management Layer**

• **Functionality Added or Changed**
Heatwave Release Notes

Heatwave Network Layer

- Heatwave network layer optimizations have improved scalability and network performance.

Heatwave Data Management Layer

- Data loaded into Heatwave, including propagated changes, are now persisted to OCI Object Storage for recovery in case of a Heatwave node or cluster failure. Previously, data was recovered from the MySQL DB System. Loading data from OCI Object Storage is faster because data does not need to be converted to the Heatwave storage format, as is required when loading data from the MySQL DB System. If data recovery from OCI Object Storage fails, Heatwave falls back to recovering data from the MySQL DB System. Data removed from Heatwave when a table is unloaded is removed from OCI Object Storage in a background operation. For related information, see Heatwave Cluster Failure and Recovery.

Functionality Added or Changed

- Heatwave now supports `COUNT(NULL)`, except in cases where it is used as an input argument for non-aggregate operators. (Bug #33005146)

- Full support was added for the `DISTINCT` modifier. Previously, multiple instances of `(DISTINCT value)` expressions in a query were only permitted if the same `value` was specified. (Bug #32865043, Bug #33007714)

- Heatwave now supports the `WITH ROLLUP` modifier in `GROUP BY` clauses.

- Heatwave now supports window functions. For optimal performance, window functions in Heatwave utilize a massively parallel, partitioning-based algorithm. For more information, see Window Functions.

Changes in Heatwave 8.0.26 (2021-07-23, General Availability)

- Advisor
- Auto Parallel Load
- Auto Scheduling
- Functionality Added or Changed

Advisor

- The new Heatwave Advisor provides string column encoding and data placement key recommendations based on machine learning models, data analysis, and Heatwave query history. Implementing Heatwave Advisor recommendations can improve query performance and reduce the amount of memory required on Heatwave nodes.

  The Heatwave Advisor also provides a Query Insights feature, which provides runtimes for successfully executed queries, and runtime estimates for EXPLAIN queries, queries cancelled using Ctrl+C, and queries that fail due to out of memory errors. Runtime data is useful for query optimization, troubleshooting, and estimating the cost of running a particular query or workload.

  The Heatwave Advisor is implemented as a stored procedure named `heatwave_advisor`, which resides in the MySQL `sys` schema. Running Advisor involves issuing a `CALL` statement for the stored procedure with optional arguments.

  ```sql
  CALL sys.heatwave_advisor (options);
  ```
Heatwave Release Notes

For more information about the HeatWave Advisor, see Workload Optimization using Advisor.

Auto Parallel Load

- The new HeatWave Auto Parallel Load utility automates the process of preparing and loading tables into HeatWave and loads data using an optimized number of parallel load threads.

The HeatWave Auto Parallel Load utility is implemented as a stored procedure named heatwave_load, which resides in the MySQL sys schema. Running Auto Parallel Load involves issuing a CALL statement for the stored procedure, which takes a list of schemas and options as arguments.

```sql
CALL sys.heatwave_load (db_list, [options]);
```

For more information about the HeatWave Auto Parallel Load utility, see Loading Data Using Auto Parallel Load.

Auto Scheduling

- The HeatWave query scheduling algorithm was improved. The revised algorithm prioritizes queries based on estimated cost and wait time in the queue, which enables dynamic, workload-aware query prioritization. Previously, queries were prioritized using a static cost-based prioritization model.

Functionality Added or Changed

- `DATE_ADD()` and `DATE_SUB()` functions now support precision INTERVAL values (DECIMAL, DOUBLE, and FLOAT). (Bug #32725985, Bug #32438123)

- Support was added for multiple instances of `COUNT(DISTINCT)` in a query. (Bug #32422984)

- Query compilation and processing was improved to permit combining aggregate operators into a single task in the physical query plan, which avoids fully materializing intermediate result sets. This enhancement reduces memory allocation and deallocation operations, memory usage, and execution time for affected queries.

- The cost model that estimates HeatWave query runtimes can now use statistics from previously executed queries, which improves the accuracy of query runtime estimates.

- HeatWave now supports `CREATE TABLE ... SELECT` statements where the `SELECT` query is offloaded to HeatWave and the table is created on the MySQL Database Service instance. This feature improves `CREATE TABLE ... SELECT` performance in cases where the `SELECT` portion of the statement is a long running, complex query. For more information, see CREATE TABLE ... SELECT Statements.

- Support was added for `REGEXP_REPLACE()` and `REGEXP_SUBSTR()` regular expression functions, and error messaging was improved for `REGEXP()` function syntax mismatches, expression errors, and input argument errors.

Changes in HeatWave 8.0.25 (2021-05-11, General Availability)

Functionality Added or Changed

- Support was added for `CAST()` of ENUM column values to CHAR or VARCHAR where the ENUM value is cast to a FLOAT value, as in the following example:

```sql
SELECT CAST(CAST(enum_col AS FLOAT) AS CHAR(3)) FROM tbl_name;
```

(Bug #32618454)
• Support was added for `SELECT DISTINCT` queries that order the result set by a column that is not defined in the `SELECT` list. For example, the following query can now be offloaded to HeatWave for execution:

```
SELECT DISTINCT a FROM t1 ORDER BY c DESC;
```

(Bug #32583856)

• Query plan statistics are now collected and stored in a statistics cache when a query is executed in HeatWave. When a new query shares query execution plan nodes with previously executed queries, the actual statistics collected from previously executed queries are used instead of estimated statistics, which improves query execution plans, cost estimations, execution times, and memory efficiency.

The statistics cache is an LRU structure. When cache capacity is reached, the least recently used entries are evicted from the cache as new entries are added. The maximum number of entries permitted in the statistics cache is defined by the `rapid_stats_cache_max_entries` setting. The number of entries permitted by default is 65536, which is enough to store statistics for 4000 to 5000 unique queries of medium complexity.

• Support was added for:
  - `CAST() AS YEAR`. Both variable-length and dictionary-encoded string columns values are supported.
  - The `FORMAT()` function. Variable-length-encoded string columns are not supported.

**Changes in HeatWave 8.0.24 (2021-04-20, General Availability)**

**Functionality Added or Changed**

• Comparison of different temporal type values is now supported. For example, a query that compares `DATE` values to `TIMESTAMP` values can now be offloaded to HeatWave. (Bug #32420986)

• Range operators on `VARLEN`-encoded string columns are now supported. For example, the following query, where `L_LINESTATUS` is a `VARLEN`-encoded string column, can now be offloaded to HeatWave:

```
SELECT COUNT(*) FROM lineitem WHERE L_LINESTATUS  >= 1 and L_LINESTATUS <= 10;
```

(Bug #31721399)

• HeatWave now supports `INSERT ... SELECT` statements where the `SELECT` query is offloaded to HeatWave and the result set is inserted into a table on the MySQL Database Service instance. This feature improves `INSERT ... SELECT` performance in cases where the `SELECT` portion of the statement is a long running, complex query. For more information, see INSERT ... SELECT Statements.

• `VARLEN`-encoded columns are now supported as data placement keys. For information about the data placement feature, see Defining Data Placement Keys.

• Failure handling was improved for queries involving unsupported internal data types. Such queries now exit with an error indicating that the internal data type of the query is not supported.

**Changes in HeatWave 8.0.23-u2 (2021-03-15, General Availability)**

**Functionality Added or Changed**

• Support was added for the following aggregate functions:
  - `STD()`
• STDDEV()
• STDDEV_POP()
• STDDEV_SAMP()
• VAR_POP()
• VAR_SAMP()
• VARIANCE()

See Aggregate Functions.

• HeatWave now uses a priority-based scheduling mechanism based on query cost estimates to schedule queries for execution. Previously, queries were executed in the order of arrival. The scheduling mechanism prioritizes short running queries over long running queries to reduce overall query execution wait times. For more information, see Auto Scheduling.

Changes in HeatWave 8.0.23-u1 (2021-02-09, General Availability)

Functionality Added or Changed

• String column encoding support was added for TEXT-type columns. See Encoding String Columns.

• UNION and UNION ALL support was extended. The clauses are now supported at any location in a query that is permitted by MySQL.

• The following date and time functions are now supported:
  • TO_SECONDS()
  • UNIX_TIMESTAMP()
  • FROM_UNIXTIME()
  • TIME_TO_SEC()

See Date and Time Functions.

The following date and time functions are now supported with VARLEN-encoded columns:

• TO_DAYS()
• DAYOFYEAR()
• QUARTER()
• TO_SECONDS()

See Date and Time Functions.

The following string functions are now supported with VARLEN-encoded columns:

• ORD()
• ASCII()
See String Functions and Operators.

```sql
SET timezone = timezone with the timezone value specified as an offset from UTC in the form of [H]H:MM and prefixed with a + or - is now supported with the UNIX_TIMESTAMP() and FROM_UNIXTIME() functions.
```

- Offset is now supported with the LIMIT clause:
  ```sql
  SELECT * FROM tbl LIMIT offset, row_count;
  ```
  The PostgreSQL syntax is also supported:
  ```sql
  SELECT * FROM tbl LIMIT row_count OFFSET offset;
  ```

- New Performance Schema tables provide access to query and execution statistics:
  - `performance_schema.rpd_exec_stats`
  - `performance_schema.rpd_query_stats`

Changes to HeatWave Performance Schema tables:

- An NDV (Number of Distinct Values) column was added to the `performance_schema.rpd_columns` table.
- A ROWS column that shows the total number of rows in a table was added to the `performance_schema.rpd_tables` table.
- A MEMORY_USAGE column that shows node memory usage was added to the `performance_schema.rpd_columns` table.
- The `performance_schema.rpd_nodes` DRAM column was renamed to MEMORY_TOTAL. The MEMORY_TOTAL column shows the total memory allocated to a HeatWave node.

See HeatWave Performance Schema Tables.

Index

A
  Advisor, 6
  aggregate functions, 8
  ASCII(), 9
  Auto Parallel Load, 6

B
  bloom filter optimization, 4

C
  CAST(), 5, 7
  change propagation, 3
  column limits, 4
  comparisons, 3
  CONVERT_TZ(), 4
  COUNT(DISTINCT), 6
  COUNT(NULL), 5
CREATE TABLE ... SELECT, 6

D
data compression, 3
data placement, 8
data types, 8
DATE_ADD(), 6
DATE_SUB(), 6
DAYOFYEAR(), 9
DISTINCT, 5

F
FORMAT(), 7
FROM_UNIXTIME(), 9

G
GROUP BY, 5

I
IN(), 3
INSERT ... SELECT, 8

L
LAST_DAY, 4
LIMIT, 9

M
machine learning, 3
MAKEDATE(), 5

N
networking, 5

O
Object Storage, 4
OFFSET, 9
ORD(), 9

P
Performance Schema, 4
planned shutdown, 4

Q
QUARTER(), 9
queries, 5
query cost model, 6
query scheduling, 6, 8

R
rapid_stats_cache_max_entries, 7
recovery, 4, 5
REGEXP_REPLACE(), 6
REGEXP_SUBSTR(), 6
rpdcolums table, 9
rpdd_exec_stats table, 9
rpdd_nodes table, 9
rpdd_query_stats, 4
rpdd_query_stats table, 9

S
scalability, 5
SELECT DISTINCT, 7
SET timezone, 9
statistics, 7
string column encoding, 8

T
temporal type comparison, 8
TEXT, 9
TIME_TO_SEC(), 9
TO_DAYS(), 9
TO SECONDS(), 9

U
UNION, 9
UNION ALL, 9
UNIX_TIMESTAMP(), 9

V
variable-length encoding, 8
VARLEN encoding, 9
views, 4

W
WEEK(), 5
window functions, 5
WITH ROLLUP, 5

Y
YEARWEEK(), 5