Simple Replay Workload Analysis

Replay ID: 2023-03-11T00:56:48.015366+00:00_ra3-redshift-cluster-testing_3fefa

Cluster ID	Start Time	End Time	Instance Type	Nodes	Replay Tag
ra3-redshift-cluster-testing	2023-03-11T00:56:48+00:00	2023-03-11T01:26:45+00:00	ra3.xlplus	2	

Replay Summary

- Attempted to replay 62 queries, 62 transactions, 12 connections.
- Successfully replayed 0 out of 62 (0%) transactions.
- Successfully replayed 12 out of 62 (19%) queries.
- Encountered 0 connection errors and 0 transaction errors
- Replay finished in 0:29:57.022777.

This report summarizes the performance of the replayed workload shown above.

Glossary

The following terms are used in this report:

- **Compile Time** is the total amount of time spent compiling a query.
- **Queue Time** is the amount of time a query spends waiting before executing in a workload management (WLM) queue.
- **Execution Time** is how long a query spends in the execution phase.
- **Query Latency** is the total runtime of a query in Redshift.
- **Commit Queue Time** is the time a transaction spent waiting before entering the commit phase.
- **Commit Time** the time a transaction spent being committed.

Accessing the data

All of the performance data collected for this report is available in S3 at the following location:

s3://devsaba-sr-drill/replays/2023-03-11T00:56:48.015366+00:00_ra3-redshift-cluster-testing_3fefa/

The `raw_data` directory contains the following raw CSV files unloaded from the Redshift cluster:

- `statement_types000` Statement counts by type (e.g. SELECT, COPY, etc.)
- `query_metrics000` Query-level performance data.
- `cluster_level_metrics000` Cluster-level summary of performance data. This is used to generate the Cluster Metrics table on page 2.
- `query_distribution000` User-level summary of performance data and broken down by query execution phase. This is used to generate the latency, commit, queue, compile, and execution time tables that begin on page 3.

The `aggregated_data` directory in S3 contains CSV files of the aggregated table data used to generate this report.

Workload Notes

SimpleReplay attempts to replay the source cluster workload as faithfully as possible on the target cluster. However, the replayed workload may differ from the original workload in the following ways:

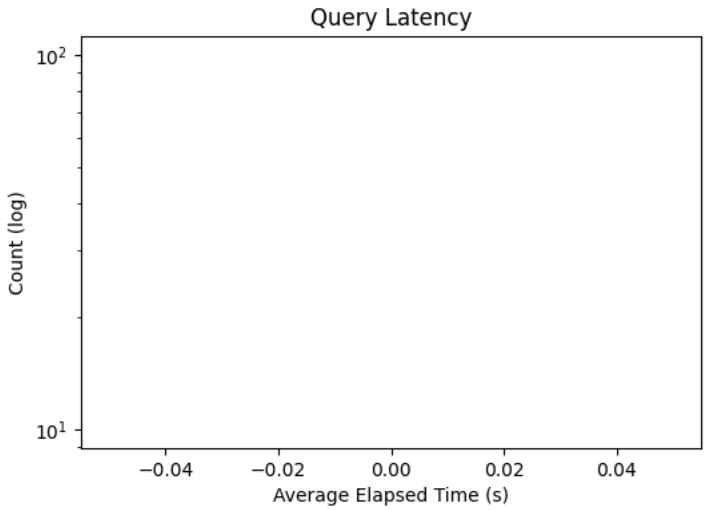
- The percentiles in this report exclude DDLs, Utility statements, and any leader node-only catalog queries.
- The reports grouped by user show the top 100 users based on the count of queries executed per user during the replay. All additional users above the top 100 are rolled up as “Others.” Data for all users is available in S3.
- Query compilation time is distributed evenly between queries that hop between service classes and can therefore occasionally result in execution or elapsed times that are less than zero for very short queries.

Query Breakdown

The table below shows the total number of queries, number of aborted queries, and number of queries executed on concurrency scaling clusters broken down by statement type.

* note that query counts are approximate and based on statement text

Statement Type	Total Count	Aborted Count
UTILITY	36	0



The histogram shows a breakdown of query latency on a log scale. The distribution show shorter running queries on the left and longer running queries on the right.

Cluster Metrics

The table below shows performance statistics broken down by cluster-level workload metric.

* note that query latency excludes compile time

Measure	Avg(s)	Std Dev(s)	P25(s)	P50(s)	P75(s)	P99(s)
Query Latency	0.0	0.0	0.0	0.0	0.0	0.0
Queue Time	0.0	0.0	0.0	0.0	0.0	0.0
Execution Time	0.0	0.0	0.0	0.0	0.0	0.0

Query Latency

Query latency is the combined amount of time a query spends queued in WLM and executing. Note that this does not include query compilation time.

User	Query Count	Avg(s)	Std Dev(s)	P25(s)	P50(s)	P75(s)	P99(s)
awsuser	36	0.0	0.0	0.0	0.0	0.0	0.0

Compile Time

Redshift compiles queries before executing them, and then caches the compiled result. This table shows how much time is spent compiling queries, broken down by user. Note that a workload run on a new cluster may have higher compile time than the original source cluster workload since it may not benefit from prior caching.

User	Query Count	Avg(s)	Std Dev(s)	P25(s)	P50(s)	P75(s)	P99(s)

Queue Time

Queue time shows how much time a query is spent waiting to start executing. It should usually be considered together with Execution Time (since high queue time + low execution time is the same to the user as low queue time + high execution time).

User	Query Count	Avg(s)	Std Dev(s)	P25(s)	P50(s)	P75(s)	P99(s)
awsuser	36	0.0	0.0	0.0	0.0	0.0	0.0

Execution Time

Execution time shows how much time a query is spent executing. It should usually be considered together with Queue Time (since high queue time + low execution time is the same to the user as low queue time + high execution time).

User	Query Count	Avg(s)	Std Dev(s)	P25(s)	P50(s)	P75(s)	P99(s)
awsuser	36	0.0	0.0	0.0	0.0	0.0	0.0

Commit Queue Time

Transactions may be queued before before the commit phase starts. This table summarize how much time each user’s transactions spend waiting to start the commit.

User	Query Count	Avg(s)	Std Dev(s)	P25(s)	P50(s)	P75(s)	P99(s)

Commit Time

This table summarizes how much time is spent committing transactions, broken down by user.

User	Query Count	Avg(s)	Std Dev(s)	P25(s)	P50(s)	P75(s)	P99(s)