Carnegie Mellon University

Metric Collection Infrastructure

Qidu He, Dongsheng Yang, Wenxuan Qiu

Objective

- **75%: Implement a minimum set of metrics**
  - **Database level metrics: number of committed/aborted transactions**
- **100%:**
  - **Transaction level metrics: latency, number of tuple accesses/modifications**
  - **Persist to storage**
  - **Write more tests (current code coverage: 89%)**
  - ~~Push data to aggregator when a thread leaves the worker pool~~
- **112.5%/125%:**
  - **Interact with Catalog**
  - Interact with Settings Manager

Carnegie Mellon University

Testing

Six test cases in total:

1. Single thread, single statistic aggregation
2. Single thread, database level statistics aggregation
3. Single thread, database level statistics storage in SQL tables
4. Single thread, transaction level statistics aggregation
5. Single thread, transaction level statistics storage in SQL tables
6. Multiple thread, transaction level statistics storage in SQL tables

Benchmark

Two benchmarks:

Collection speed:

- single thread
- only collection, no operation
- throughput: ~ 1,000,000 event/s

Aggregation speed:

- single aggregator
- throughout: ~ 500 aggregation/s

Code Quality

Strengths:

- Flexible selection of collectable metrics
- Minimal disruption to execution paths
- Extensible to additional events, metrics, and output streams

Weaknesses:

- Stats aggregator requires a separate thread, which can potentially become a performance bottleneck
- Thread level stats collectors need to be created and destroyed with caution,

Future Work

- Memory usage

  Support memory usage collection, such as memory used per table
- CPU time

  Besides total latency, provide fine-grained timing such as CPU time per transaction
- Flexible storage

  Support more storage methods besides SQL table
- Optimization of tuple-level data

  Approximately compress tuple level metric data using randomized data structures to reduce storage space

Carnegie Mellon University

Extra Slides

Implementation

Data Holder AbstractRawData —inherit→ DatabaseMetricRawData, TransactionMetricRawData, ...

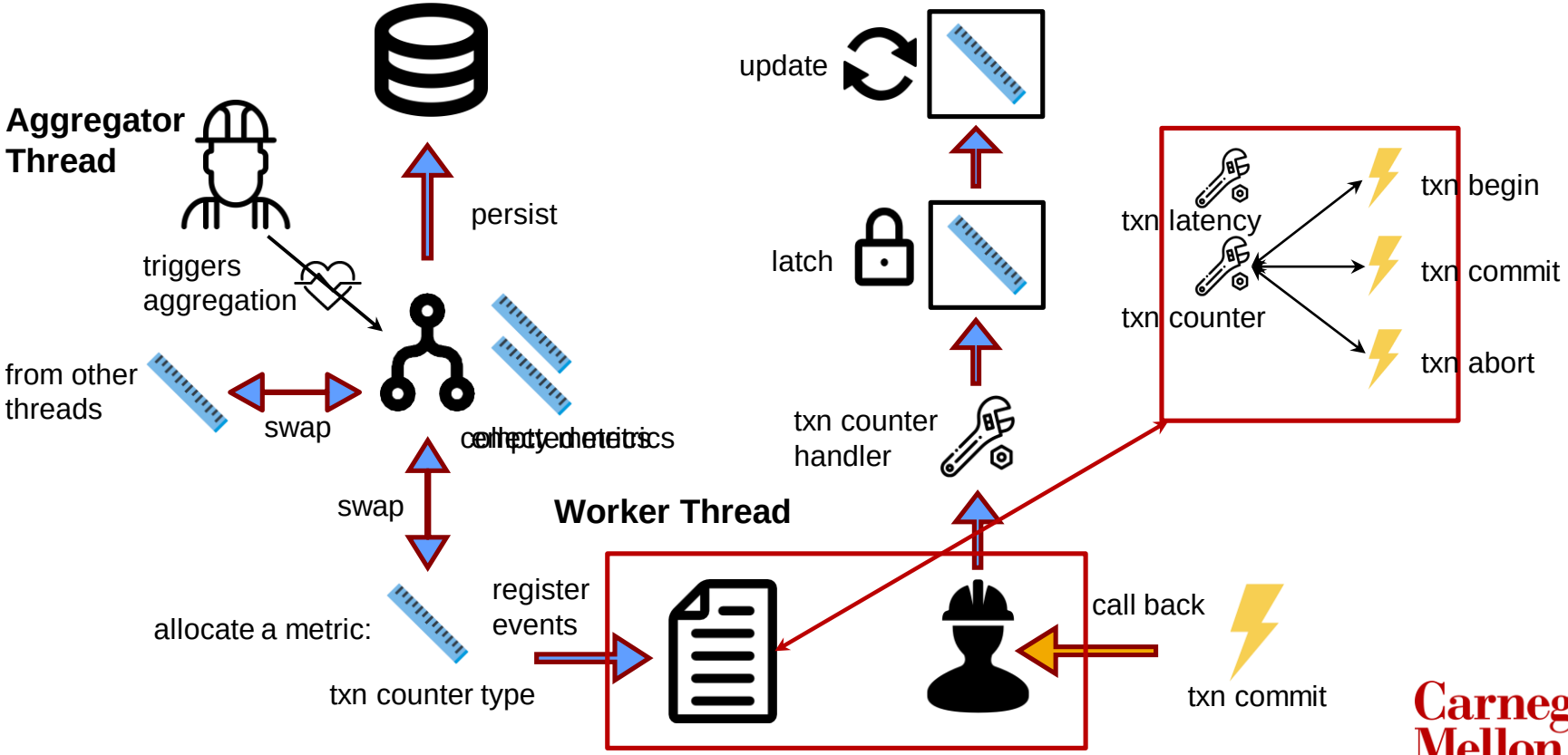
Latch Wrapper RawDataWrapper

Data Handler AbstractMetric —inherit→ DatabaseMetric, TransactionMetric, ...

Worker ThreadLevelStatsCollector

Aggregator StatsAggregator

Data Flow



Carnegie Mellon University