Workload-Aware, Self-Designing Buffers in LSM-Engines

Alexander Ott, Subhadeep Sarkar

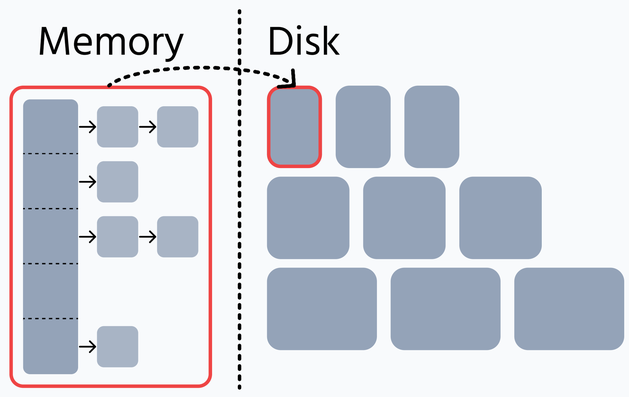
LSM Mem Buffers

LSMs Everywhere

RocksDB, SCYLLA, APACHE HBASE, Big Table, cassandra, mongoDB

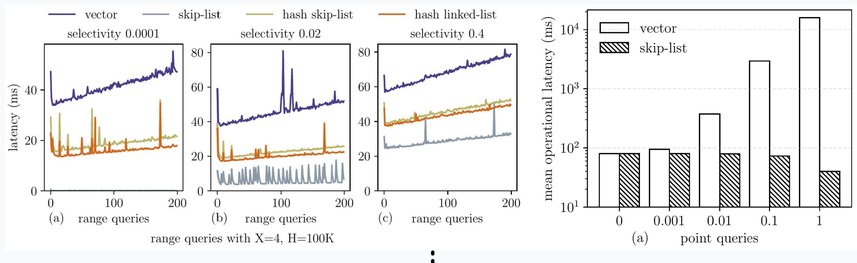
Buffered writes + Leveled storage = Fast ingestion

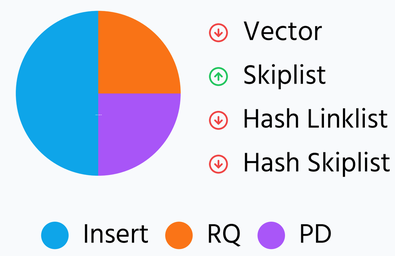
LSM Memory Buffer



- Buffers disk writes in memory
- More efficient search for recent entries

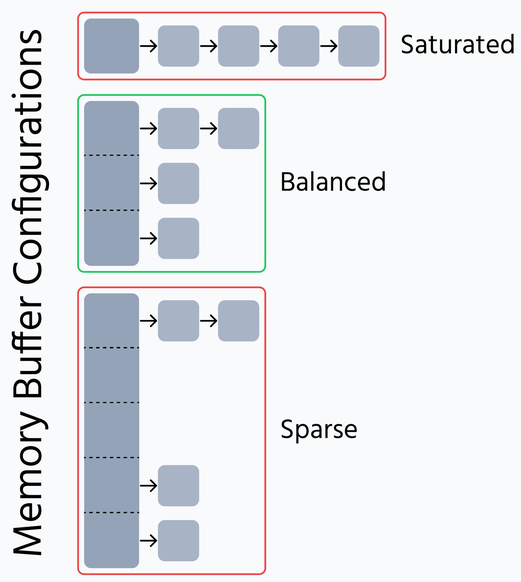
Buffer Performance

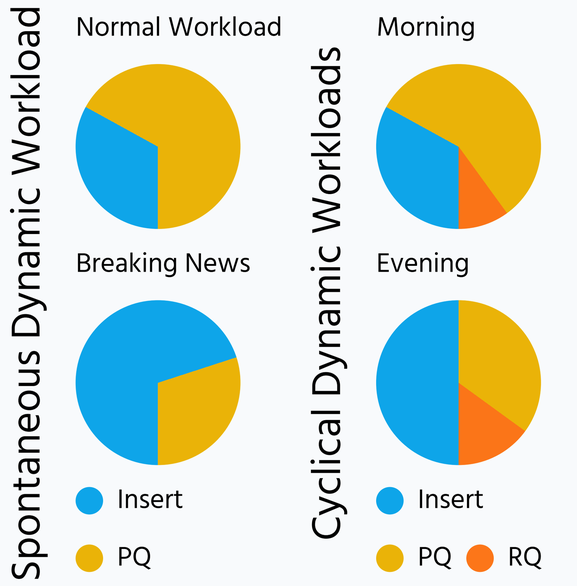


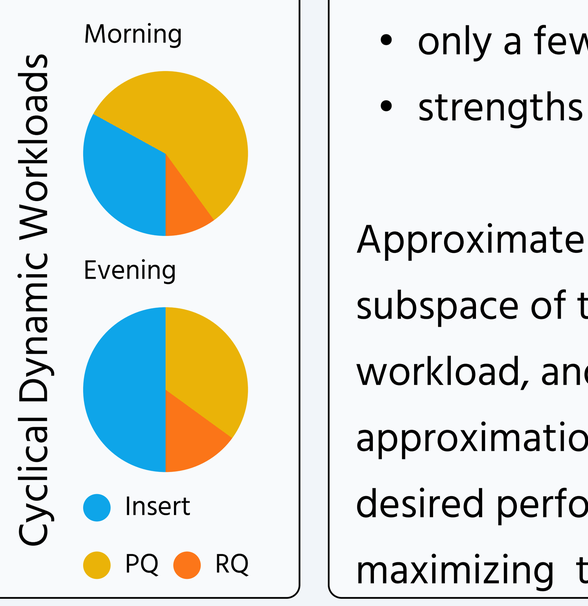


The Problem

Buffer Configuration & Dynamic Workloads







Formal Definition

- only a few memory buffer implementations to use
- strengths and weaknesses of memory buffers are not known

Approximate the function $f(C, w_{t-1}) \rightarrow c_t$ where C is a subspace of the configuration space, w_{t-1} is the previous workload, and c_t is the new configuration. Given an approximation of f, optimize $o(c_t) \rightarrow p_t$, where p_t is the desired performance metric, e.g. minimizing insert latency, maximizing throughput.

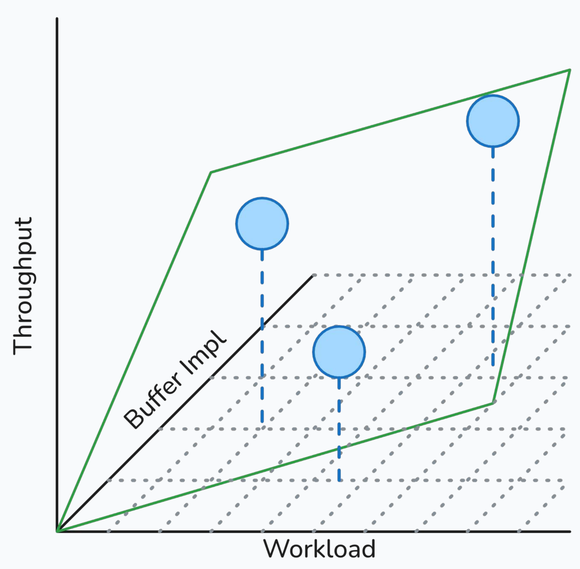
Solution

Experimentation & Implementation

- implemented a vector that is not sorted during queries
- implemented a memory buffer using the trie data structure, which is more space efficient
- dissect existing memory buffer implementations and experimentally evaluate their performance tradeoffs

- added ability to dynamically configure RocksDB at runtime
- designed an ML-model that suggests the best buffer configuration for a given workload

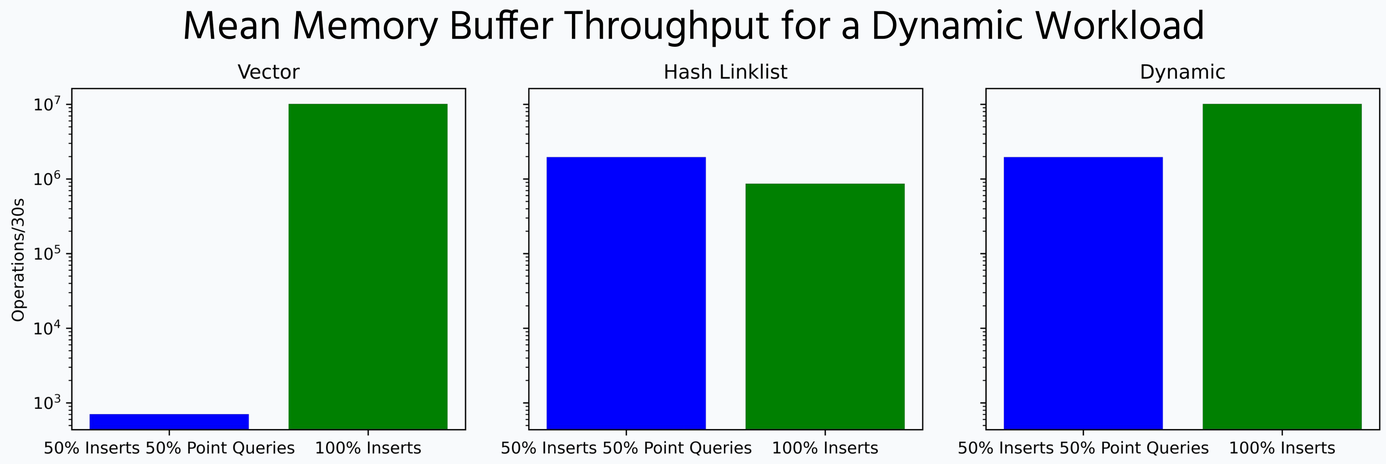
ML Optimization using TPE+CMA-ES



- Goal: optimize objective function
- Surrogate Model: Approximation of the objective function
- Acquisition function: Dictates the next location to sample, balancing between exploration and exploitation

Results

Effectiveness



- dynamic memory buffer is able to achieve high throughput in both sections of the workload

Future work

- Exploring more memory buffer implementations
  - concurrency
  - hash trie
- Introducing more knobs to optimize
  - bucket count and prefix length for hash memory buffers
- Add more optimization targets

Gordon Science Research Fellows Endowment Fund established by the Cele H. and William B. Rubin Family Fund, Inc

Brandeis UNIVERSITY
School of Arts and Sciences
Michtom School of Computer Science