Multi-Temperature LSM Tree-Based Database Storage

Hobin Yoon  
hobinyoon@gatech.edu  
Georgia Institute of Technology

Ymir Vigfusson  
ymir@mathcs.emory.edu  
Emory University & Reykjavik University

Ada Gavrilovska  
ada@cc.gatech.edu  
Georgia Institute of Technology

LSM (Log-Structured Merge) tree-based databases are popular for their high write throughput.

Cloud BigTable:  
Search, Analytics, Maps, Gmail, …

LEVELDB

RocksDB

mongoDB.

However, they can be very expensive at large scale!

Cloud BigTable:  
75,000 node Cassandra clusters, 10+ PB

NETFLIX

2,500 node Cassandra. 420TB

With traditional homogeneous storage DBs, you can get either low latency or low cost.

Meet Mutant, a multi-temperature LSM tree-based database storage, which delivers both low latency and low cost, by (a) organizing SSTables and (b) SSTable metadata by their access frequencies.

Access frequencies of the storage components

Write a record

Read a record

flush

memtable

commit log

SSTable

SSTable

SSTable

SSTables are not accessed equally

Mutant implemented by modifying RocksDB using EBS Magnetic volume $0.045 / GB / Month

RocksDB using local SSD (EC2 instance volume) $0.528 / GB / Month

Evaluation using QuizUp user profile data access traces for 16 days. 2.3 GB.

This work is supported by and CAREER #1553579.