EvenDB: Optimizing Key-Value Storage for Spatial Locality

Eran Gilad, Edward Bortnikov, Anastasia Braginsky, Yonatan Gottesman, Eshcar Hillel (Yahoo Research)
Idit Keidar (Technion), Nurit Moscovici (Outbrain), Rana Shahout (Technion)

The problem - optimizing spatially local workloads

Spatial locality
- Common with composite keys
  - field1_field2_field3
- Primary attribute has skewed dist.
- Test case (Flurry): app_timestamp, with heavy-tail app name dist. ⇒

Existing KV-stores limitations
- LSM organized by temporal locality
- Write amplification: cold ranges re-written by compactions
- All data both logged and flushed to disk

EvenDB

- Global single key caching
- One level storage
- RAM disk
- Searchable logs to delay compactions

EvenDB outperforms RocksDB when:
- Workload is spatially-local
- Or, most working set fits in RAM
- Demonstrated in real workload and synthetic YCSB benchmarks

Summary

Chunk arrangement better suited for spatially-local workloads than LSM:
- Lower write amplification
- Single level of storage
- Memory serves reads and writes

EvenDB outperforms RocksDB when:
- Workload is spatially-local
- Or, most working set fits in RAM
- Demonstrated in real workload and synthetic YCSB benchmarks