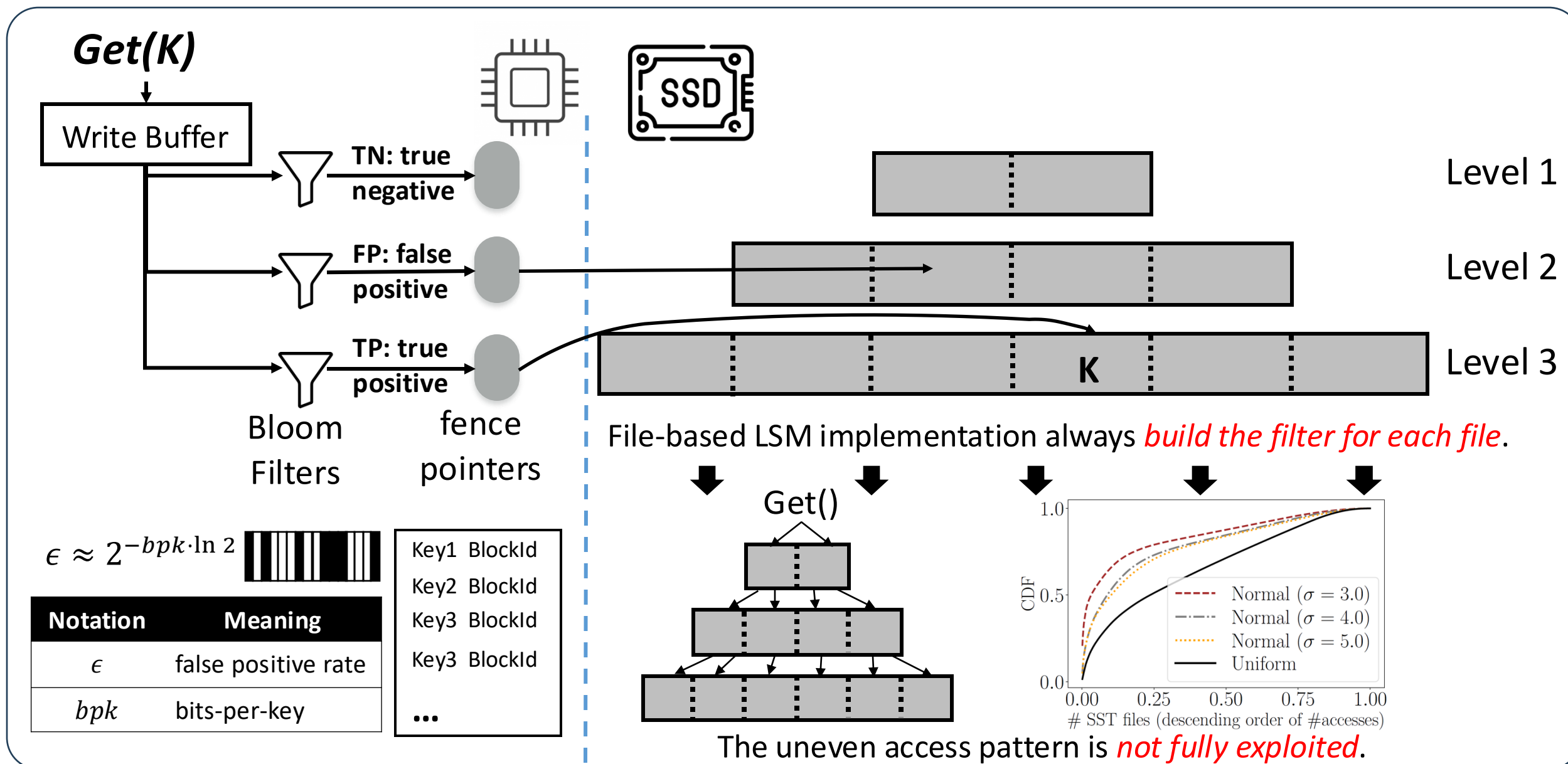


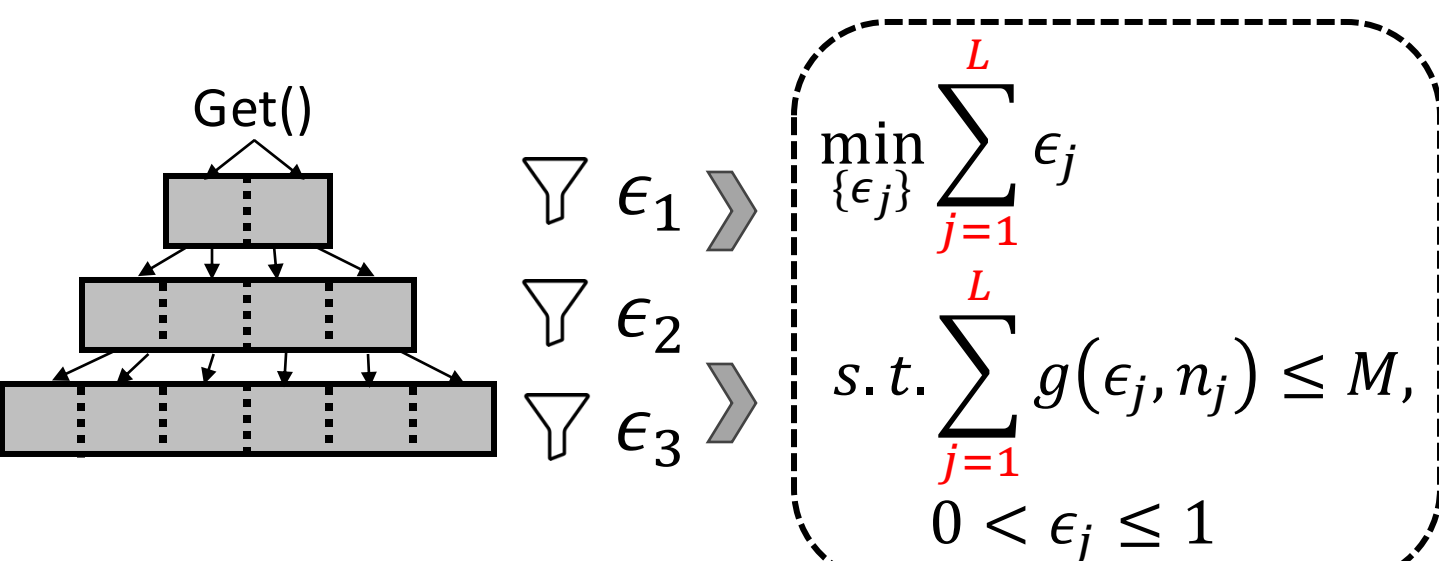
LSM-Trees are everywhere



Point Queries in LSM-Trees



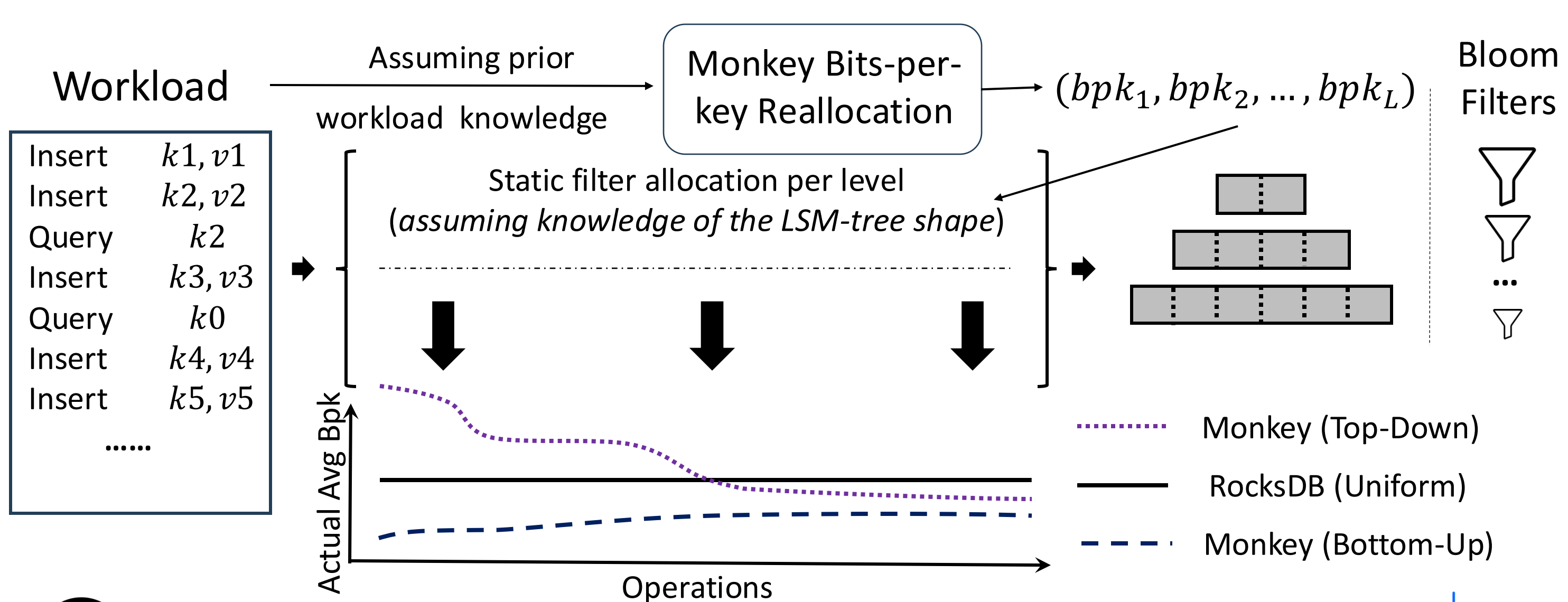
State-of-the-art: Monkey

Modeling: Each sorted run has different ϵ_j 

Notation	Meaning
ϵ_j	false positive rate in level j
L	the number of levels
n_j	the number of key-value pairs in level j
M	the total memory budget in bits
$g(\epsilon_j, n_j)$	a space function that returns the filter size in bits

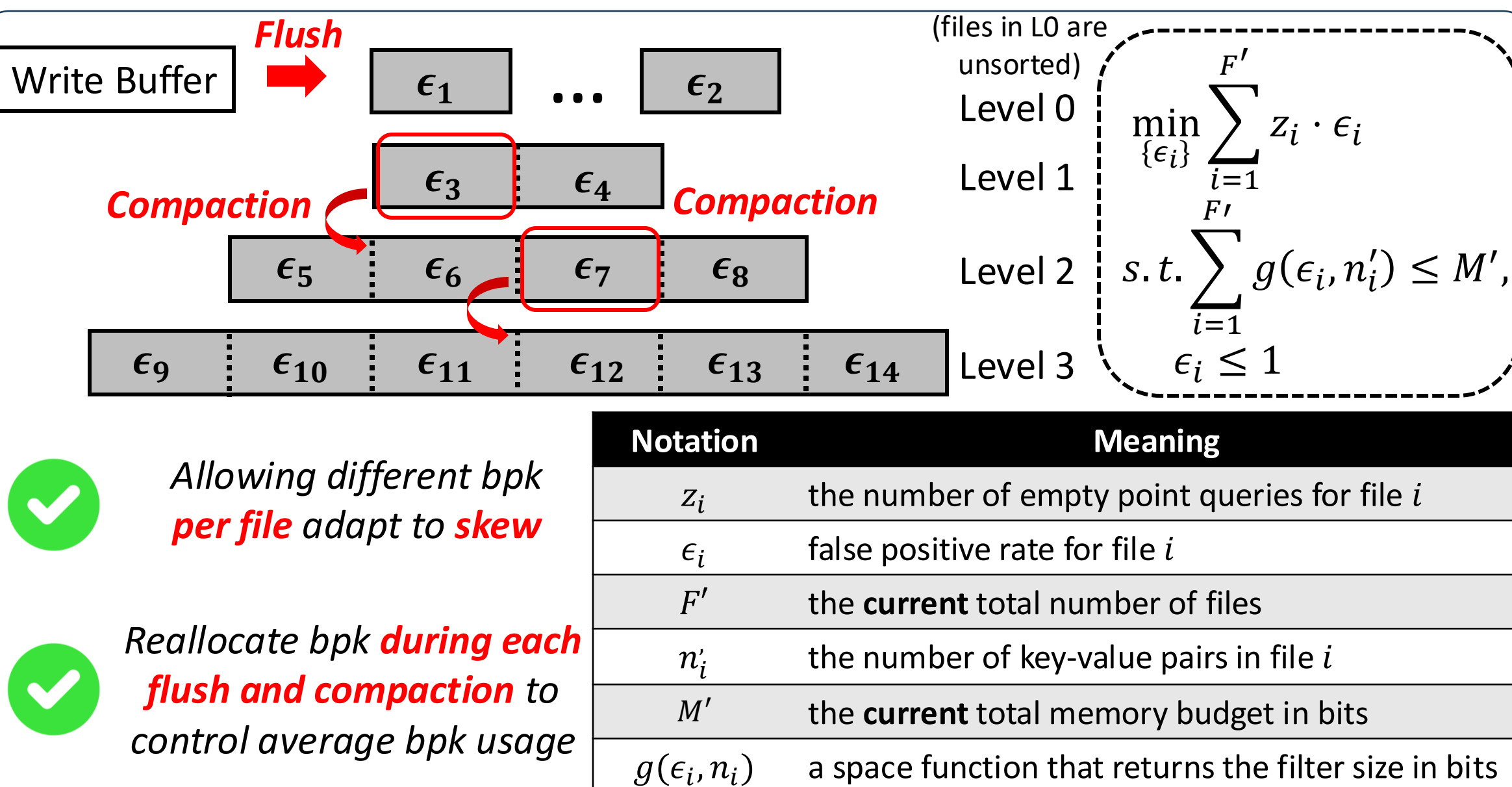
Cannot exploit *skew* query access

Workflow: Workload is known as a priori to compute bpk list

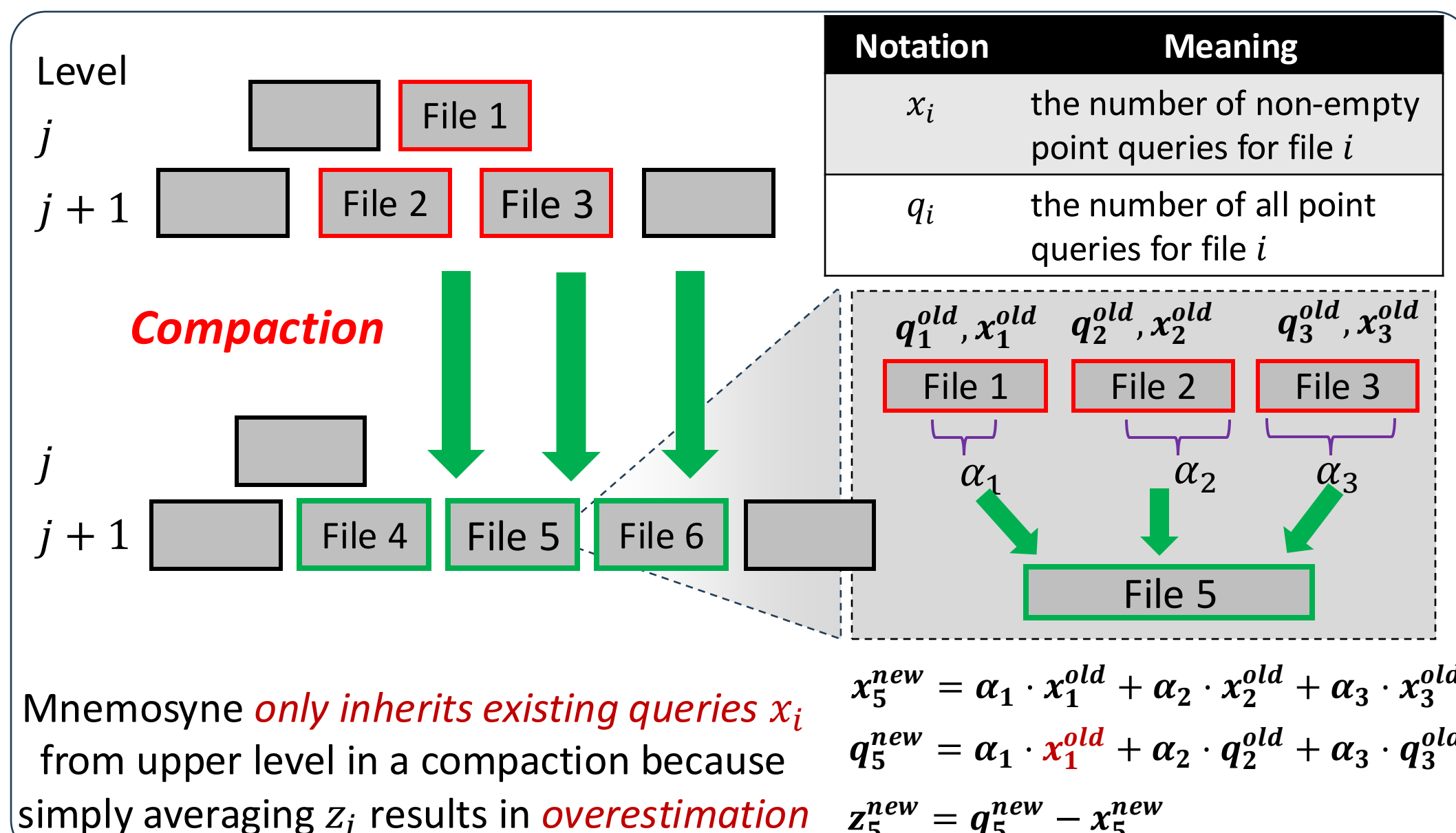


The average bits-per-key is either *overutilized or underutilized* compared to user-defined one

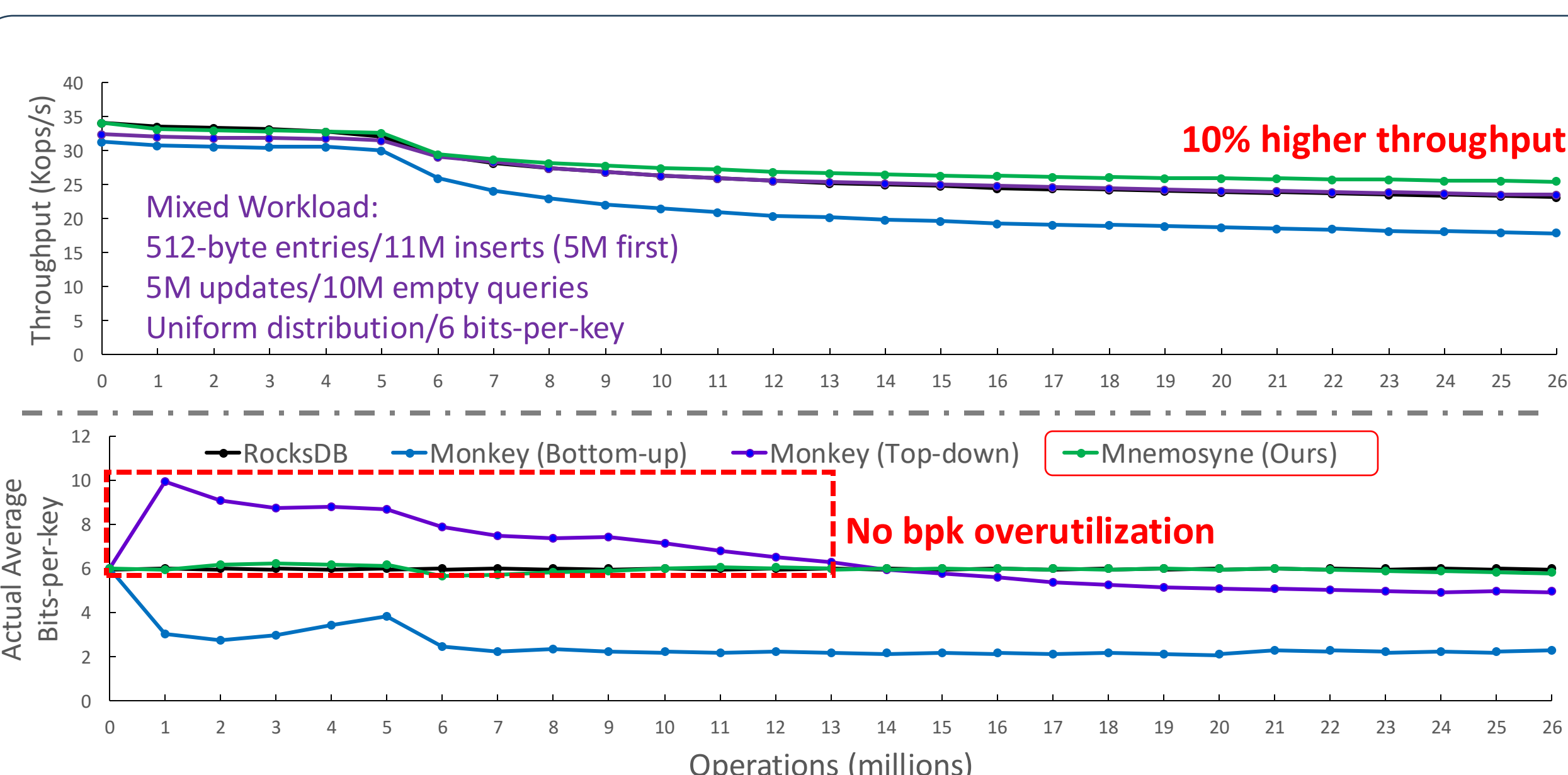
Mnemosyne: Dynamic File-wise bpk Reallocation



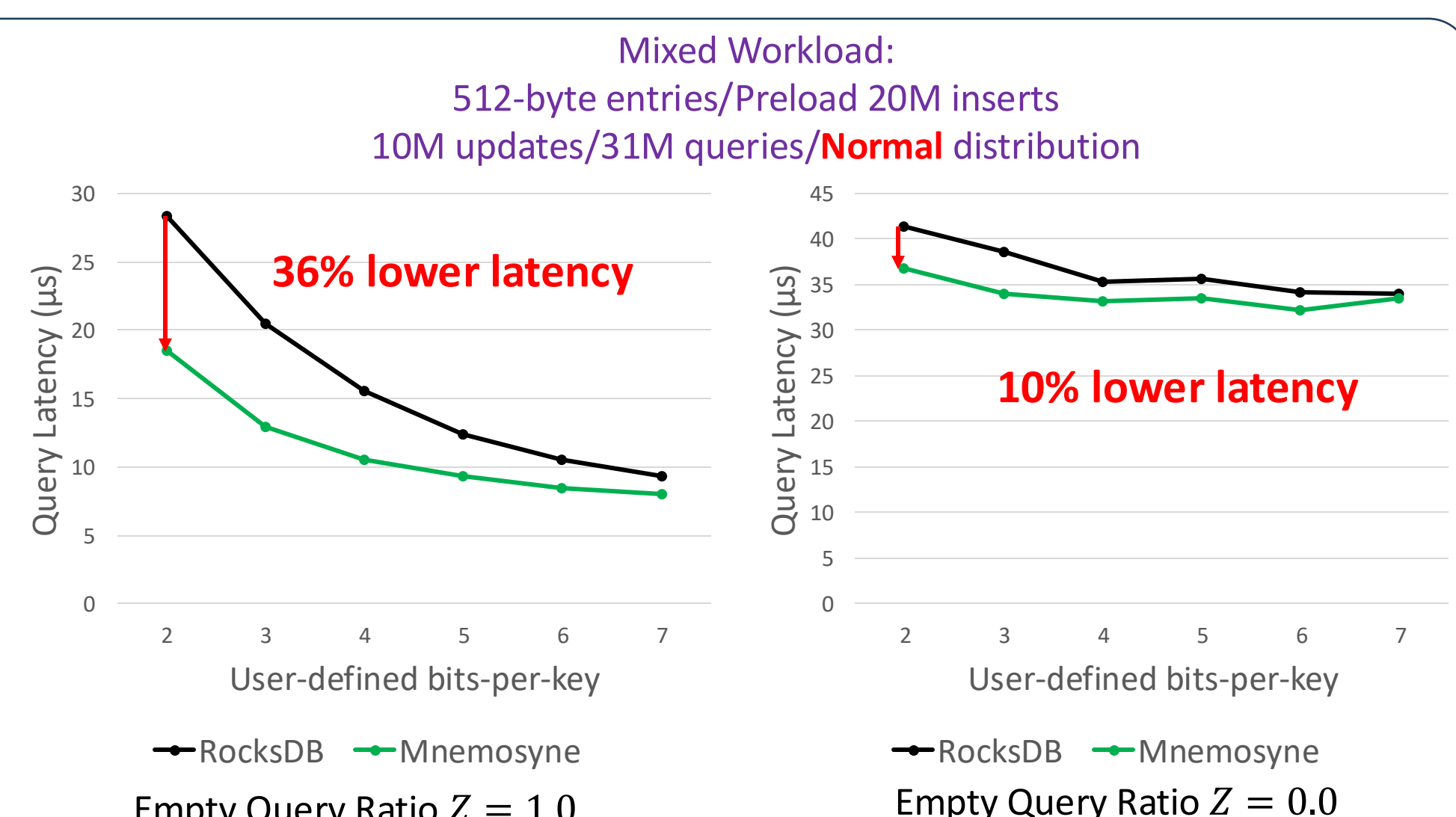
Access Stats Estimation During Compaction



Experiments



Mnemosyne has **higher throughput** than RocksDB and Monkey **without overusing** bits-per-key



Mnemosyne does **not need to have prior knowledge** and achieves much **lower query latency** for skewed accesses.

