Monkey: Optimal Navigable Key-Value Store

Niv Dayan, Manos Athanassoulis, Stratos Idreos

Buffer

key $X$

Bloom filters (M bits)

fence pointers

memory

storage

LSM-tree

merge operations

bigger filters $\rightarrow$ fewer false positives

lookup cost vs. memory

more merging $\rightarrow$ fewer runs

lookup cost vs. update cost

suboptimal Bloom filters allocation

fixed false positive rates

$\frac{p}{p}$

most memory saves at most one I/O

minimize wasted lookup I/Os

$\sum_p$ false positive rates

$= O (\log N \cdot e^{-M/N})$

optimal false positive rates

$\frac{p_0}{T^2}$

reallocating memory

write-optimized

merge greed

read-optimized

workload

hardware

optimal merge greed

0 < memory < $\infty$

buffer filters

range lookups

workload skew

caching

daslab.seas.harvard.edu/monkey