

# Benchmarking Key Value Stores with Yahoo!'s YCSB

Han Liu, Lin Xiao, Milo Polte, Julio López, Swapnil Patil, Wittawat Tantisiriroj, Garth Gibson

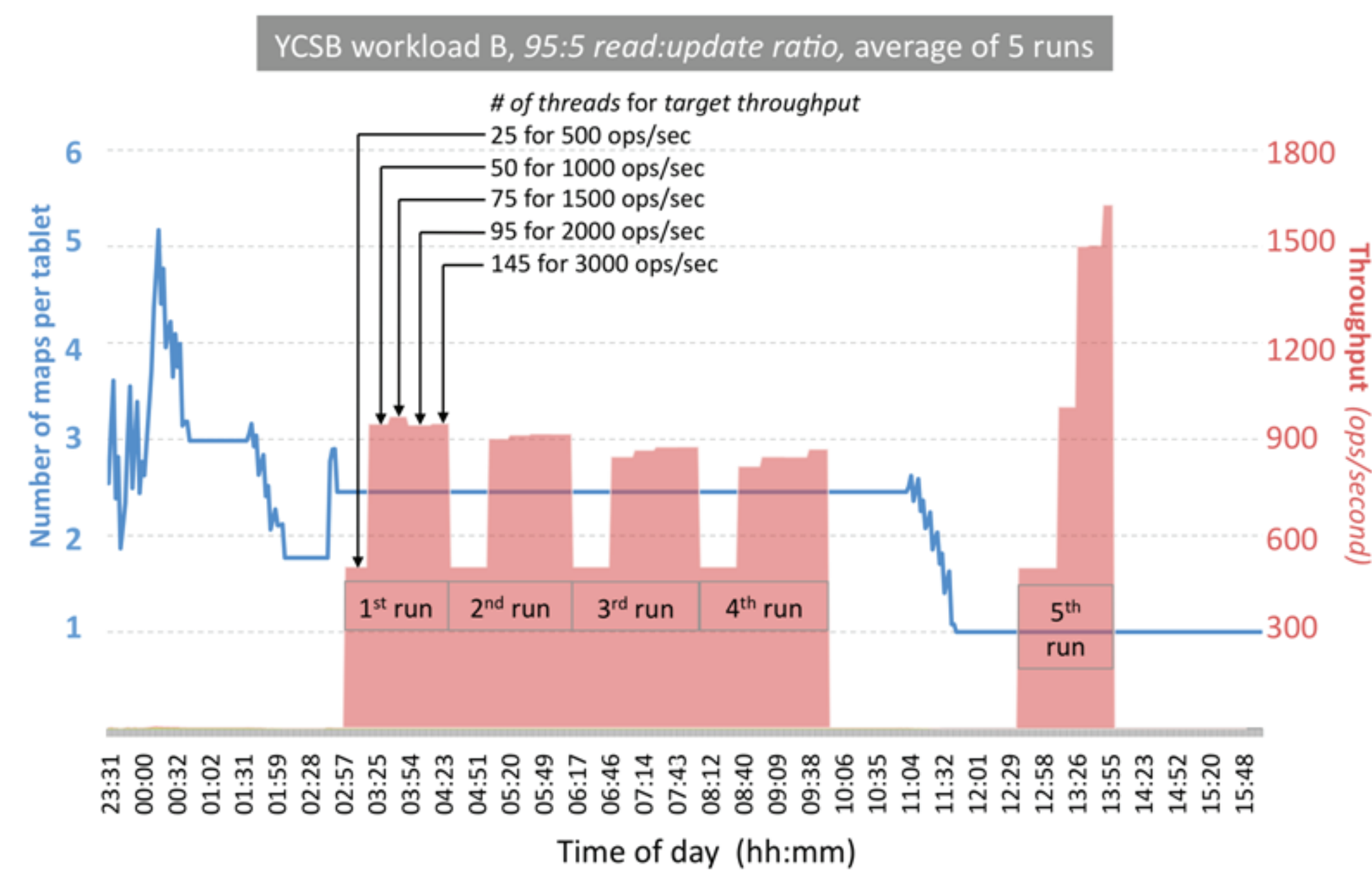
## Motivation

- Equip PDL cloud with scalable table store
- Understand and improve scalable table store with benchmarking, monitoring tools
- Explore Hbase, IBT, others

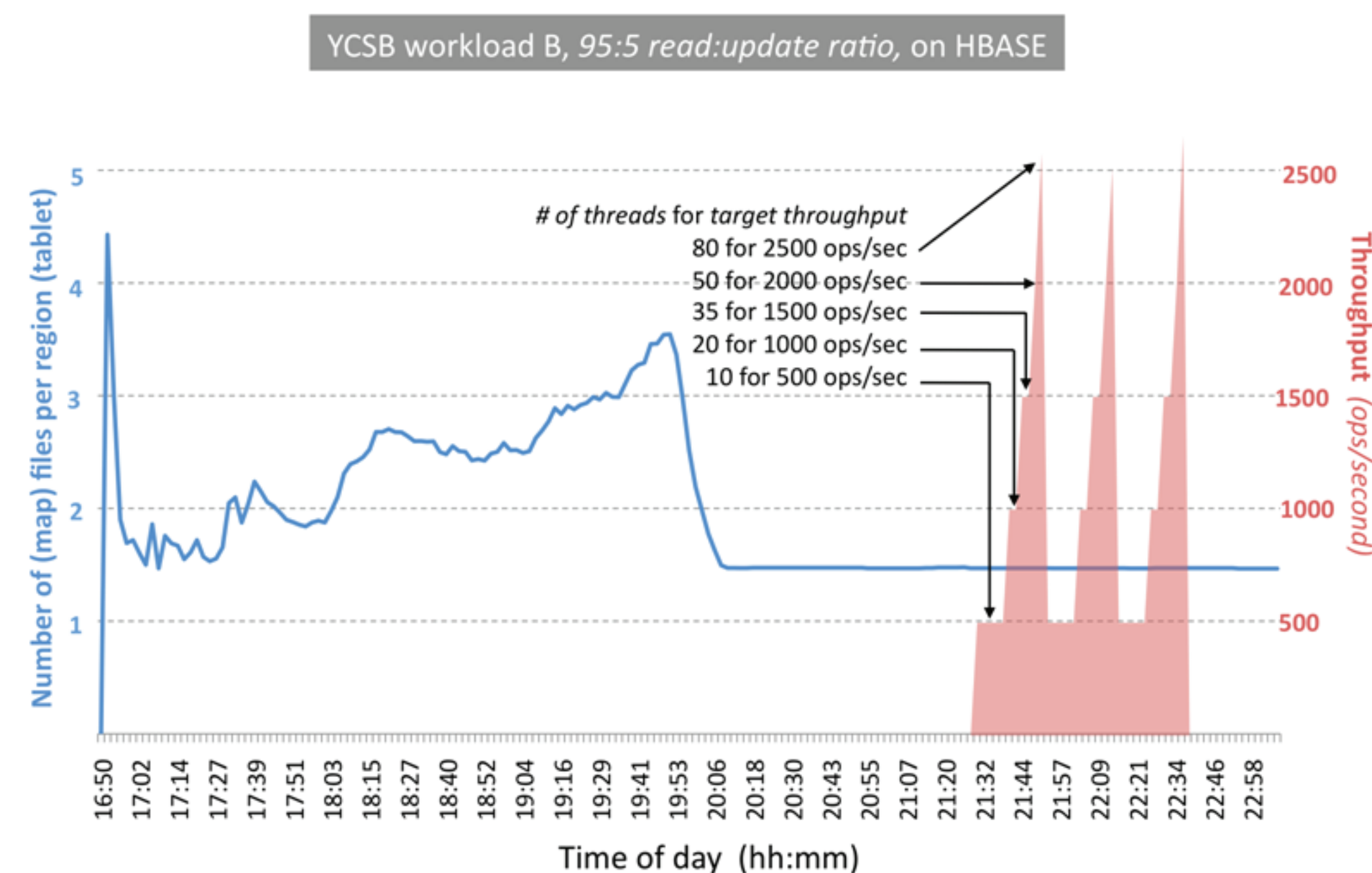
## Understanding Performance

- Visualize changes in distribution of SSTables/maps/sorted\_logs

IBT behavior



Hbase behavior



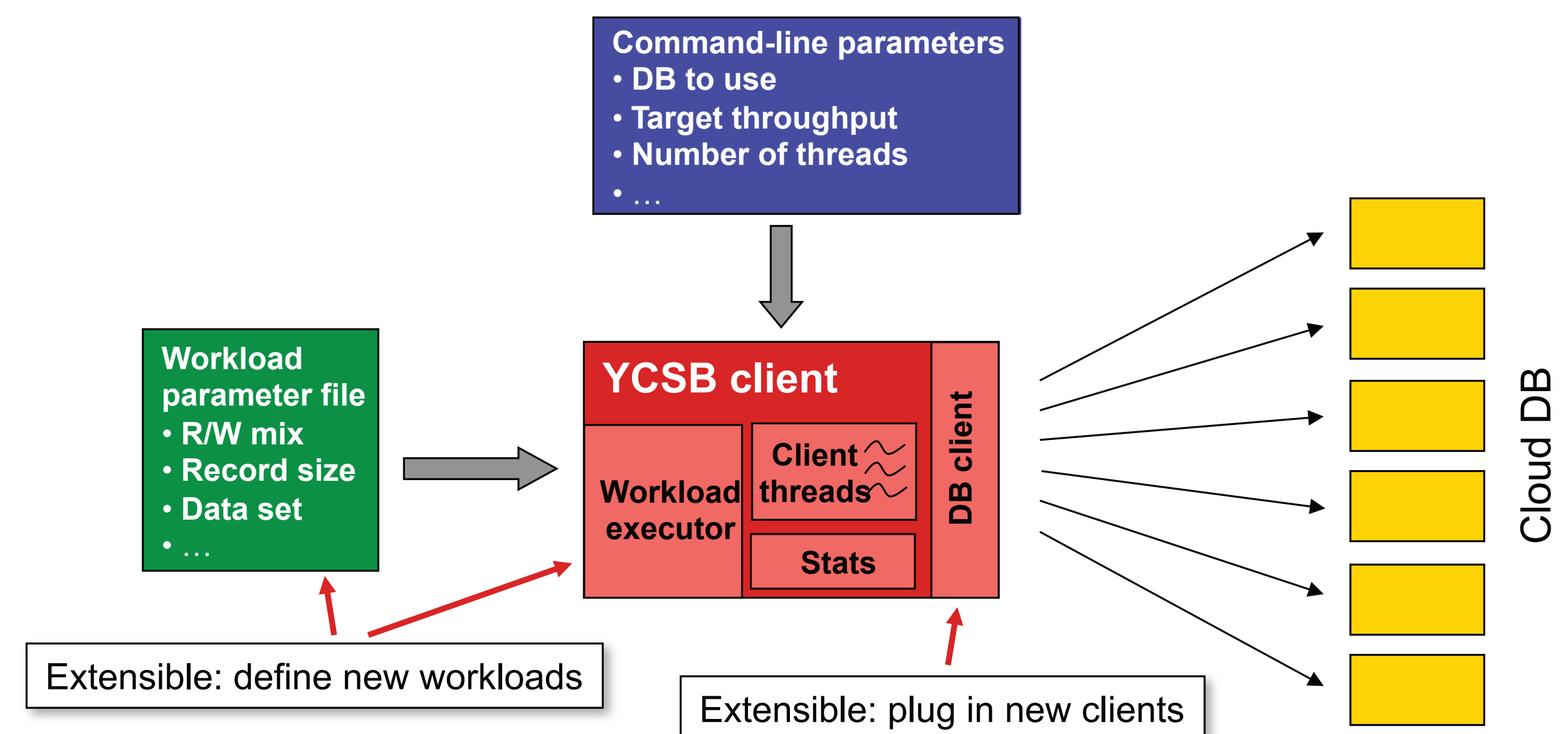
## Future Directions

- Automatic behavior charts
- Extend YCSB benchmark
  - Measure lag of insert to read
  - Automate client thread selection
- Explore bulk insert
- Load non-ordered, sparse keys
- Server side row selection

Carnegie Mellon

## Yahoo! Cloud Serving Benchmark

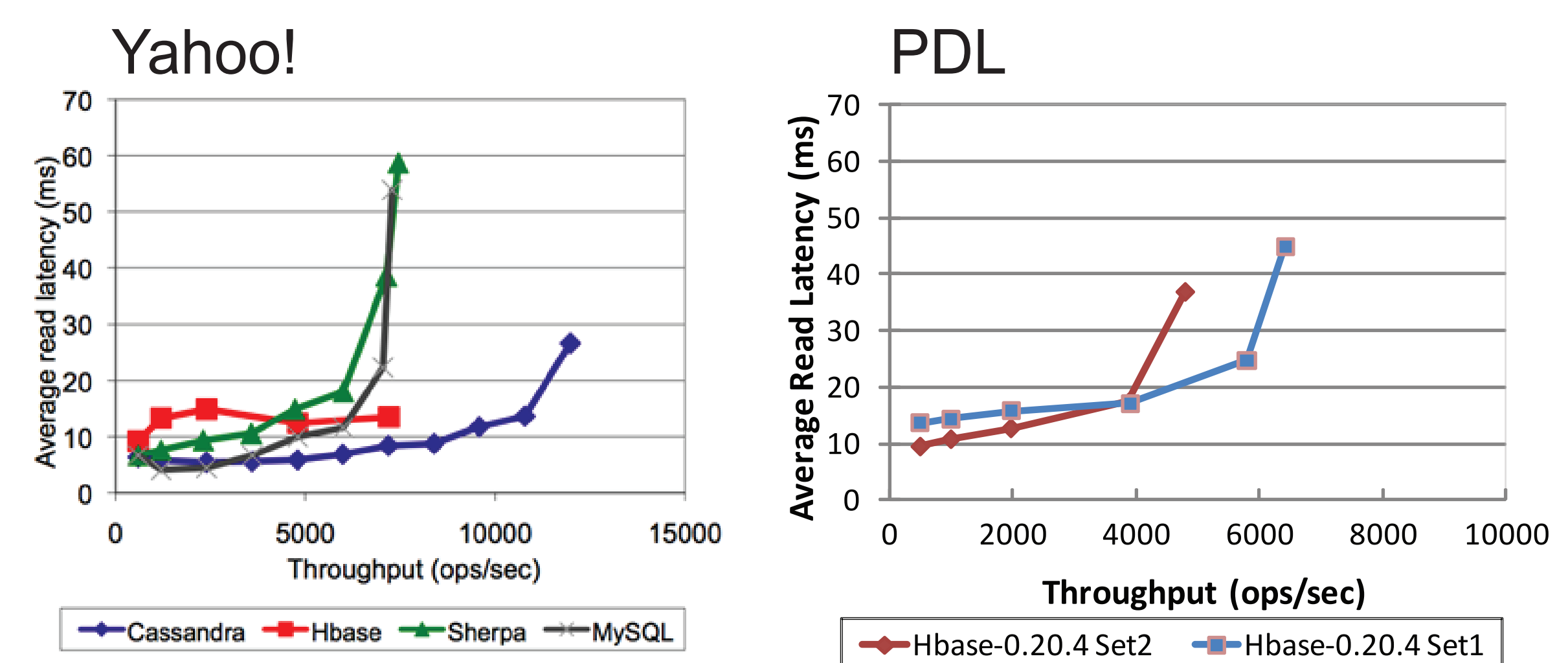
- c/o Brian F. Cooper, cooperb@yahoo-inc.com
- [wiki.github.com/brianfrankcooper/YCSB](https://github.com/brianfrankcooper/YCSB)



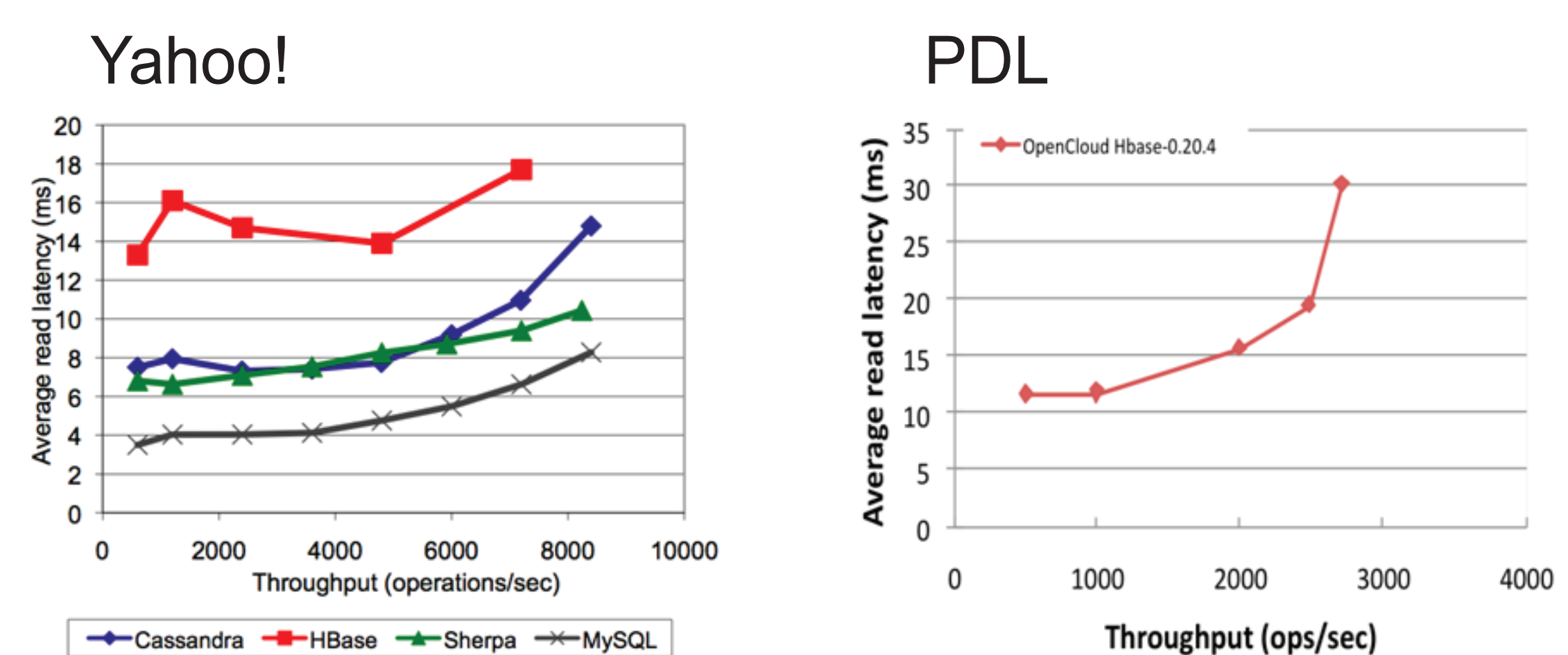
## Calibration: Yahoo! vs. PDL Hbase

- Yahoo! Hbase 0.20.3; PDL Hbase 0.20.4
- Yahoo! 6 x 15 krpm SAS; PDL 4 x 7200 rpm SATA

Workload A: 50% read 1 row, 50% update 1 cell



Workload B: 95% read 1 row, 5% update 1 cell



Workload C: 95% short (1-100) scans, 5% insert 1 row

