



Student: Tao Gao^{1,2}

Advisors: Pietro Cicotti³, Yanfei Guo⁴, Pavan Balaji⁴, Yutong Lu², Michela Tauber¹

¹University of Delaware, ²National University of Defense Technology

³San Diego Supercomputer Center, ⁴Argonne National Laboratory

Big Data on HPC

- Data-intensive applications are increasingly important on HPC platforms.
- Emerging Intel Knights Landing (KNL) has not been fully evaluated for data-intensive applications.
- Research questions: *Is KNL suitable for data-intensive applications? What is the best way to use KNL for data-intensive applications?*

Emerging KNL clusters



Evaluation Methods

- What benchmarks and datasets for our analysis?

WordCount (WC): single-pass MapReduce application.

Dataset: Wikipedia dataset from PUMA¹

K-mer counting (KC): a fundamental operation in genome analytics.

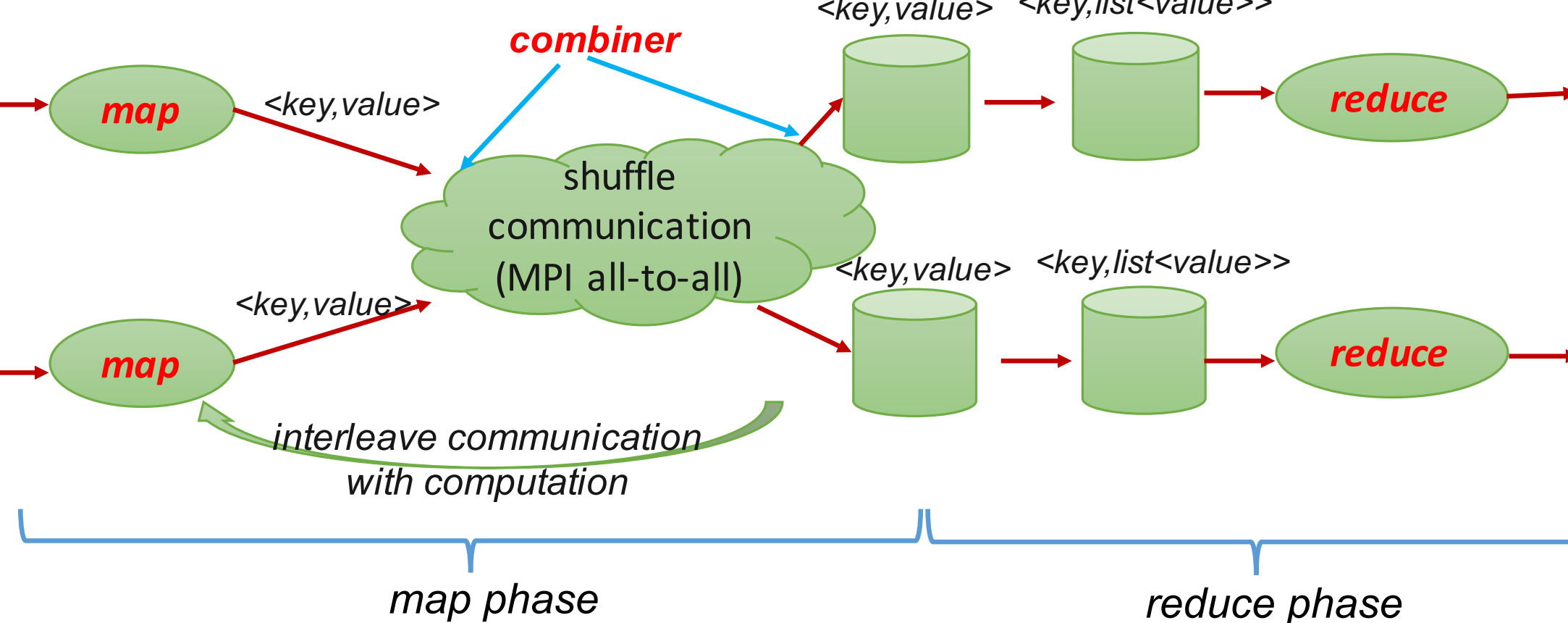
Dataset: 1000genomes dataset²

Breadth-first Search (BFS): a basic graph analysis algorithm.

Dataset: graph500³

- What programming model for the benchmarks?

A MapReduce framework (i.e. Mimir⁴) is chosen because these reasons: 1) built on top of MPI; 2) high performance and scalability; and 3) memory-efficient.



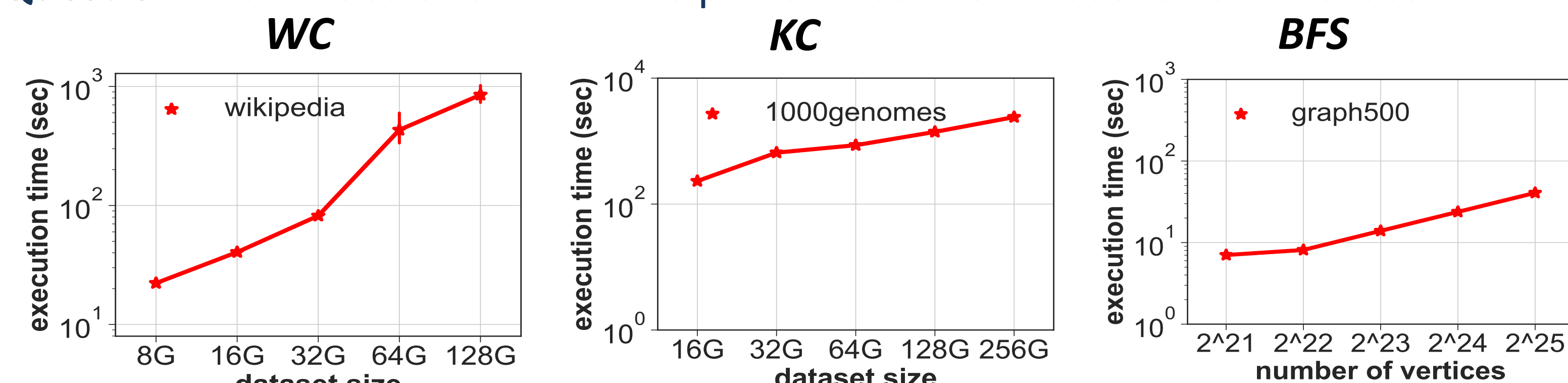
- Which KNL cluster for our tests?

The XSEDE cluster Stampede2 is used and the KNL is the Intel Xeon Phi 7250 (68 cores running at 1.4 GHz and each core supports 4 hardware threads).



Performance Analysis

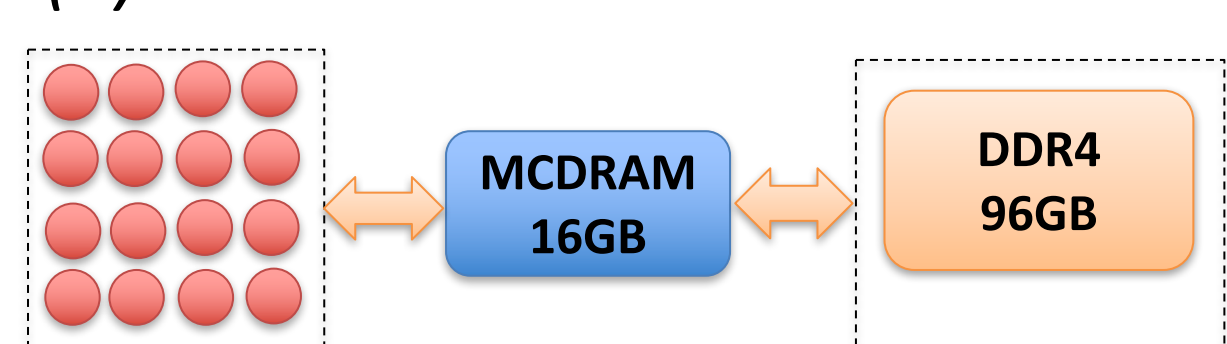
Question: How does one KNL node perform as the dataset size increases?



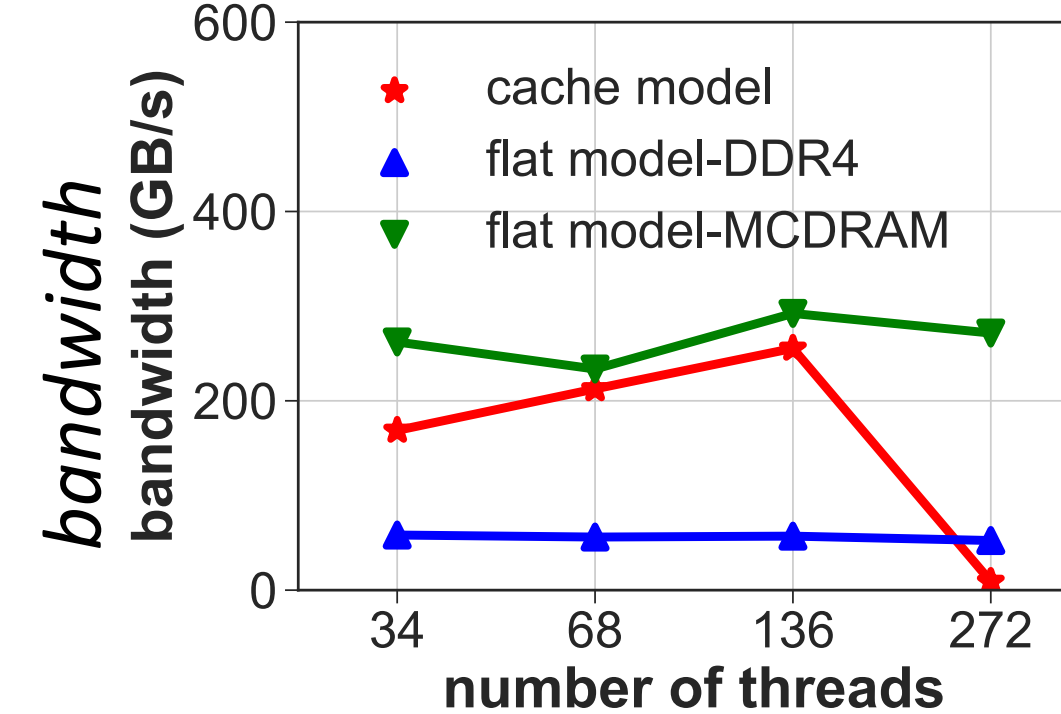
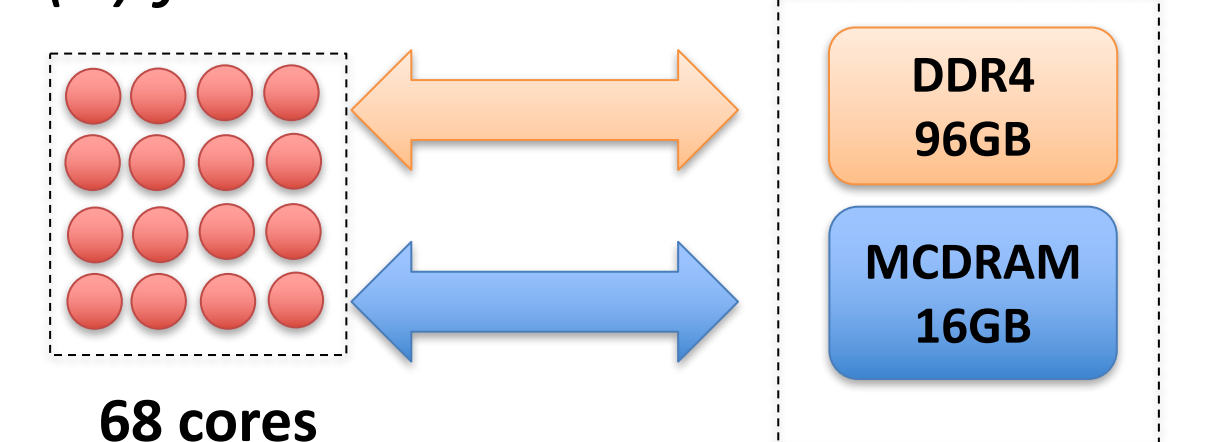
Observations: one KNL node gains linear or close to linear performance as the dataset size increases.

Question: Can data-intensive applications benefit from KNL's MCDRAM?

(1) cache model



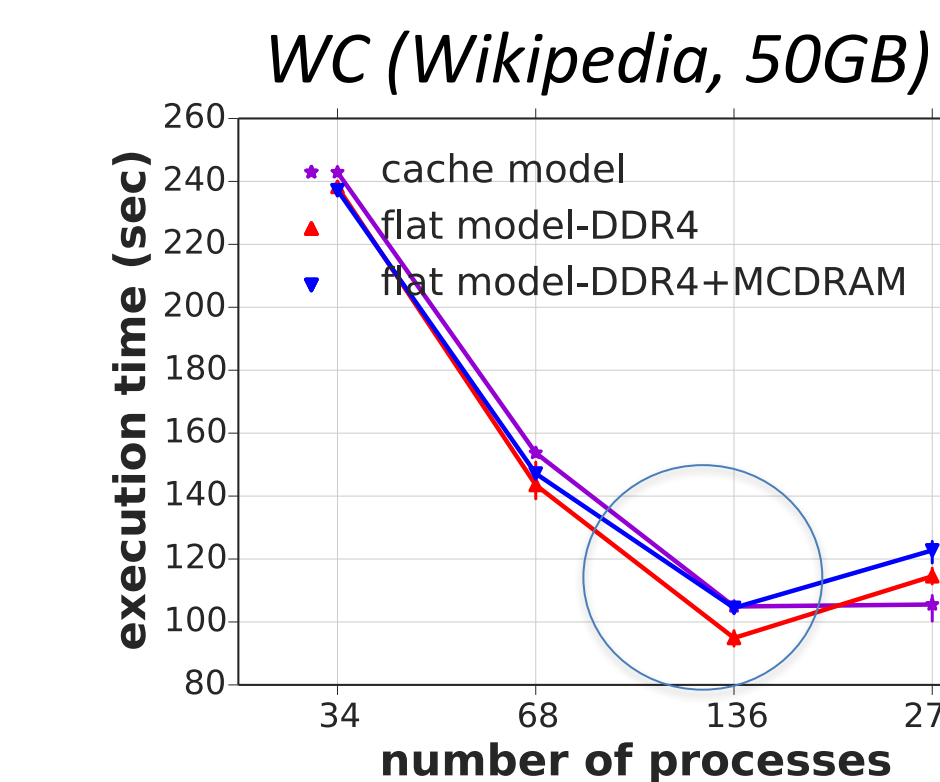
(2) flat model



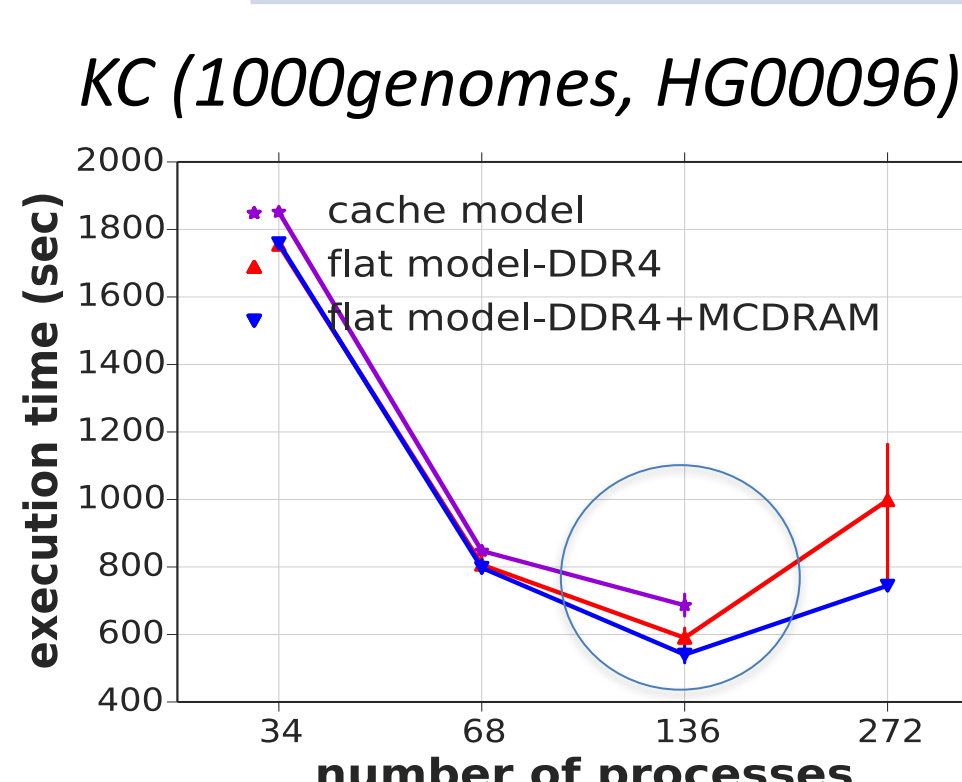
Bandwidth: flat model-MCDRAM > cache model > flat model-DDR4

Latency: cache model > flat model-MCDRAM > flat model-DDR4

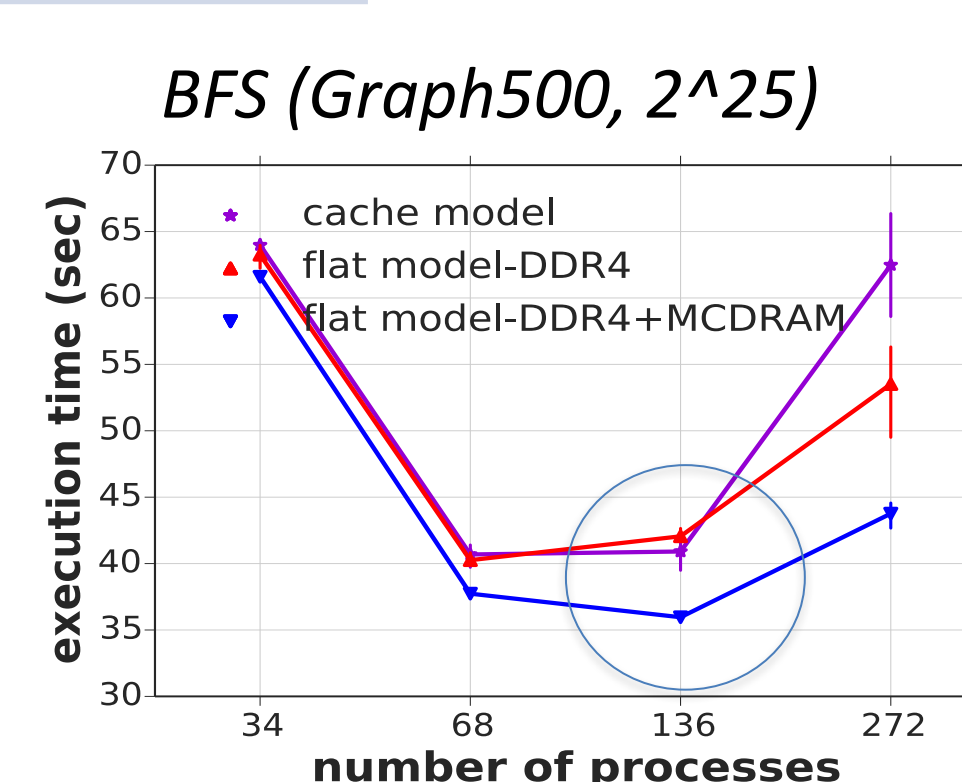
settings	latency (ns)
cache model	183.3
flat model-DDR4	149.9
flat model-MCDRAM	178.3



~20% cache miss rate
has random accesses to data exchange buffer



~10% cache miss rate
no random accesses to data exchange buffer

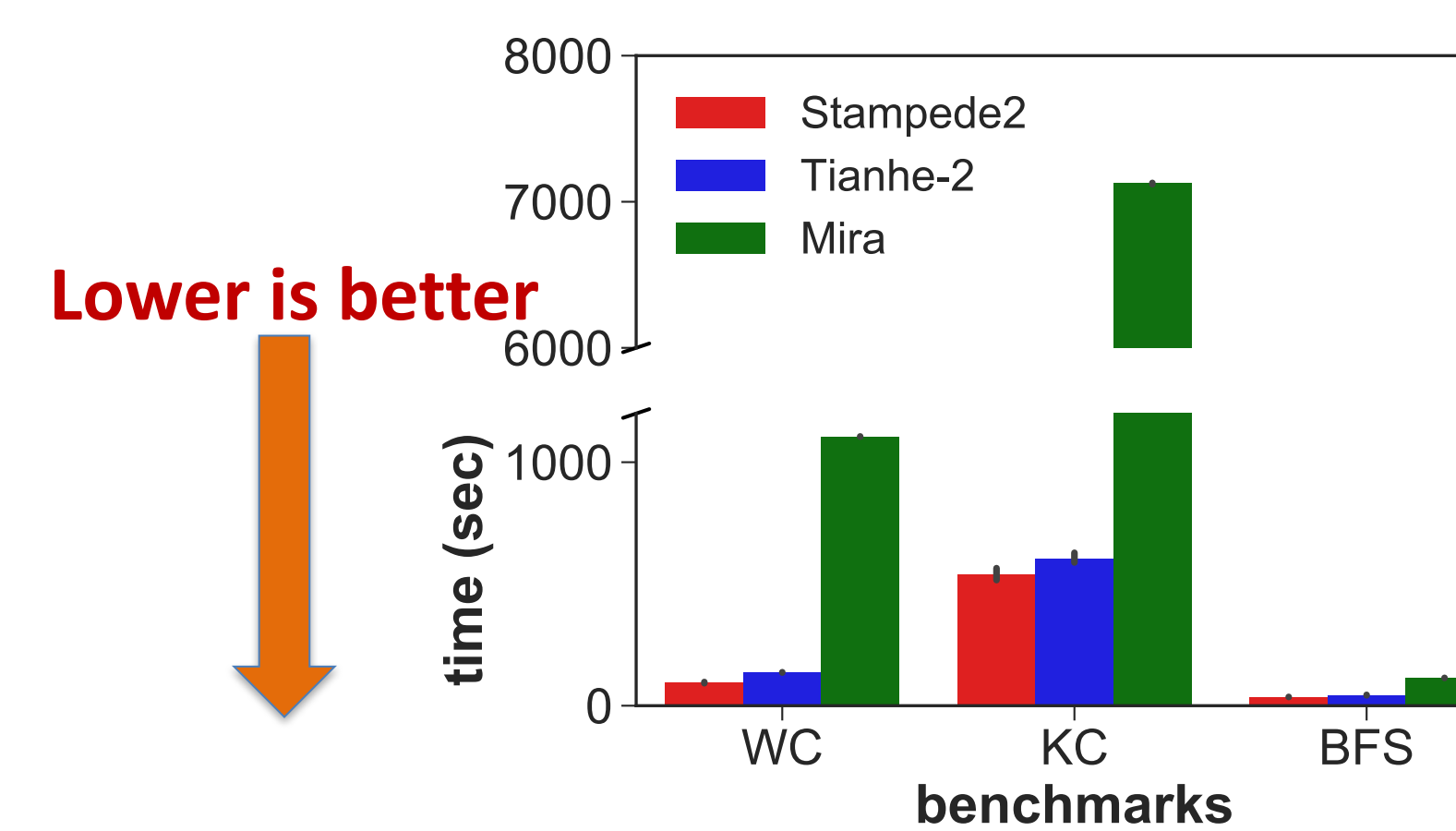


~30% cache miss rate
no random accesses to data exchange buffer

Observations: Cache model perform poorly due to high memory access latency; flat model is suitable for data-intensive applications (i.e., it provides more memory to use, its DDR has lower memory access latency in flat model); data-intensive applications benefits from MCDRAM by putting latency-bound structure on DDR and bandwidth-bound structure on MCDRAM.

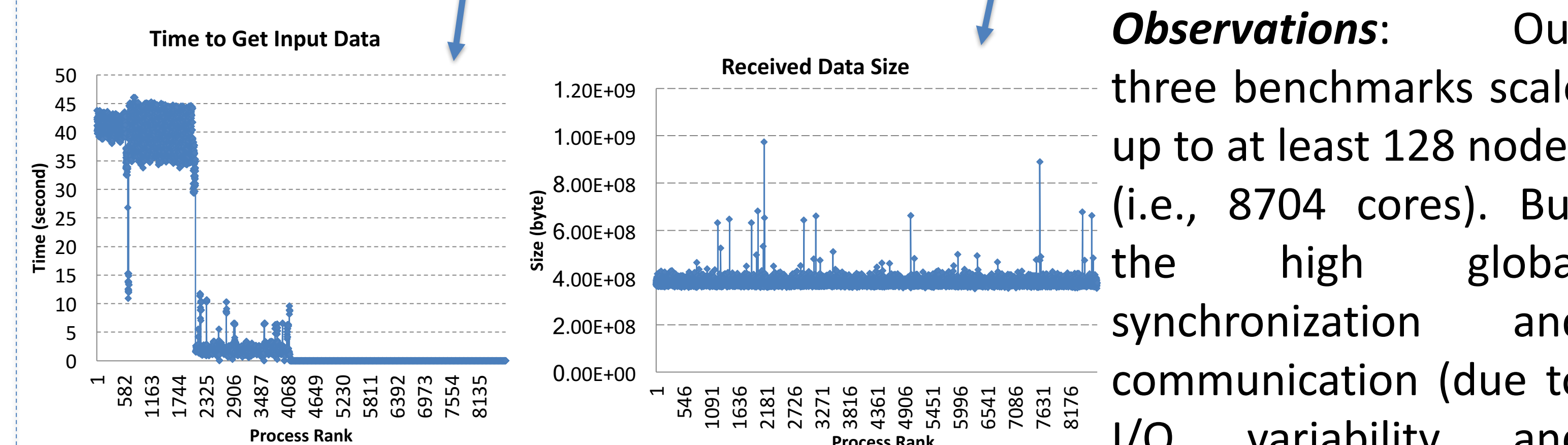
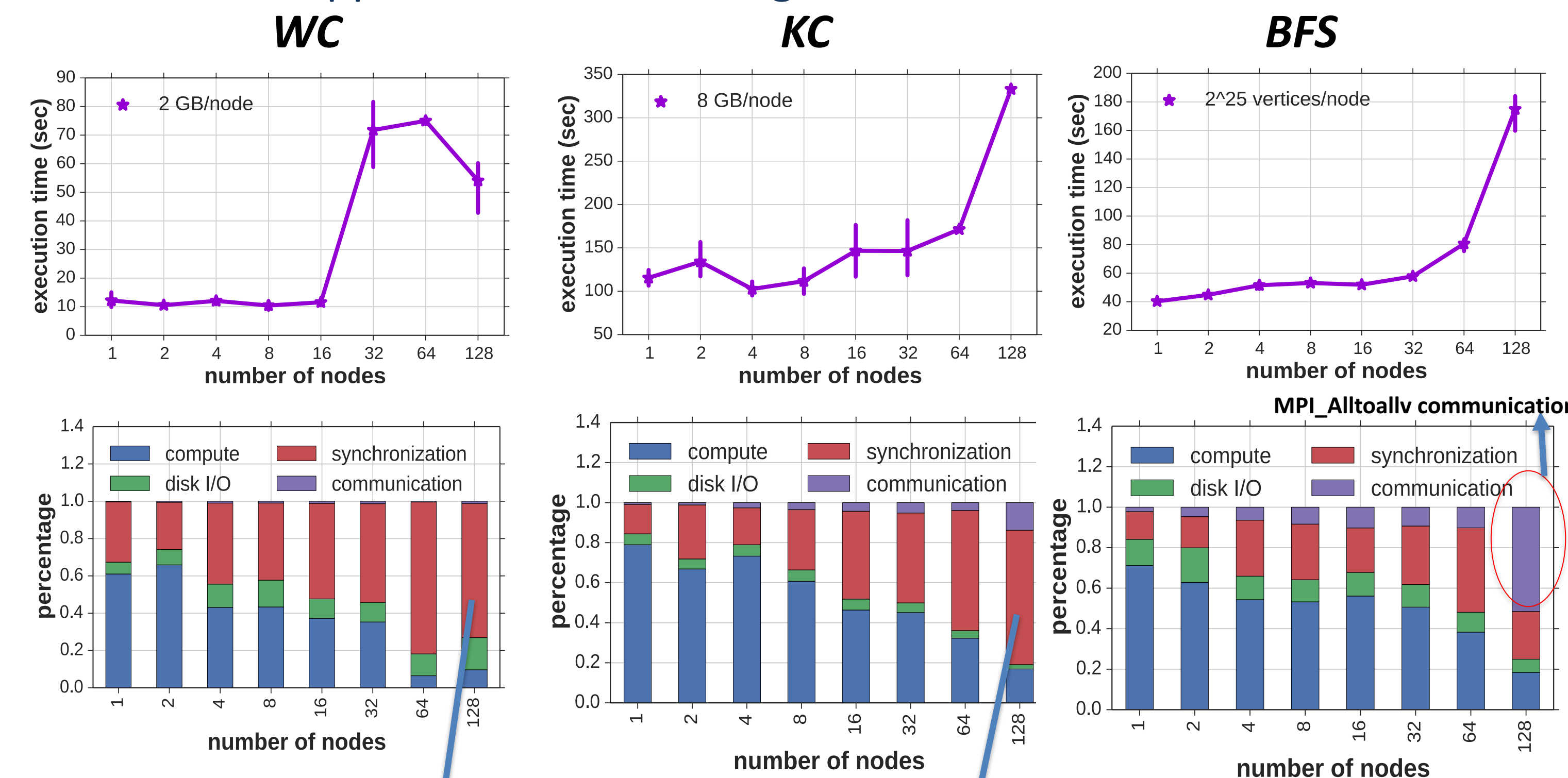
LESSONS LEARNED: KNL is a promising platform to execute data-intensive applications; applications should provide some advanced optimizations to maximize the performance: (1) make use of MCDRAM; (2) provide I/O optimizations to adapt the performance variability; (3) minimize the load imbalance in the applications; and (4) minimize global communications.

Question: How does one KNL node perform compared with other platforms?

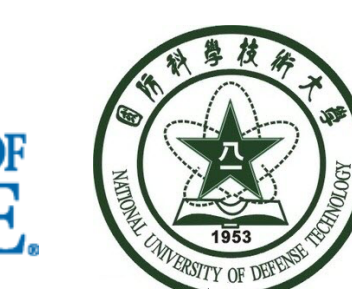


Observations: single KNL node performs better than single Intel Xeon node on Tianhe-2 (two sockets) and much better than a single IBM node on Mira (16 cores) → KNL is a promising platform for data-intensive applications.

Question: Can applications scale to large number of nodes on KNL cluster?



Observations: Our three benchmarks scale up to at least 128 nodes (i.e., 8704 cores). But the high global synchronization and communication (due to I/O variability and imbalanced load) still limits the scalability above 32 nodes.



国防科学技术大学



1 <https://engineering.purdue.edu/~puma/datasets.htm>; 2 <http://www.internationalgenome.org>; 3 <http://graph500.org> 4 T. Gao, Y. Guo, B. Zhang, P. Cicotti, Y. Lu, P. Balaji, and M. Tauber. Mimir: Memory-Efficient and Scalable MapReduce for Large Supercomputing Systems. The 31st IEEE International Parallel and Distributed Processing Symposium (IPDPS) 2017.