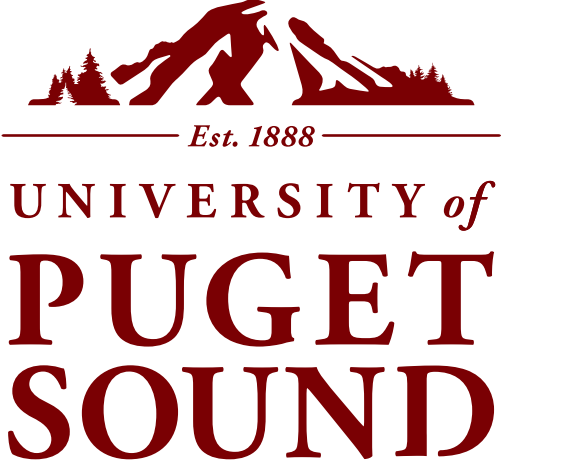


# Hadoop in Flight: Migrating Live MapReduce Jobs



Chili Johnson | David Chiu (Advisor)

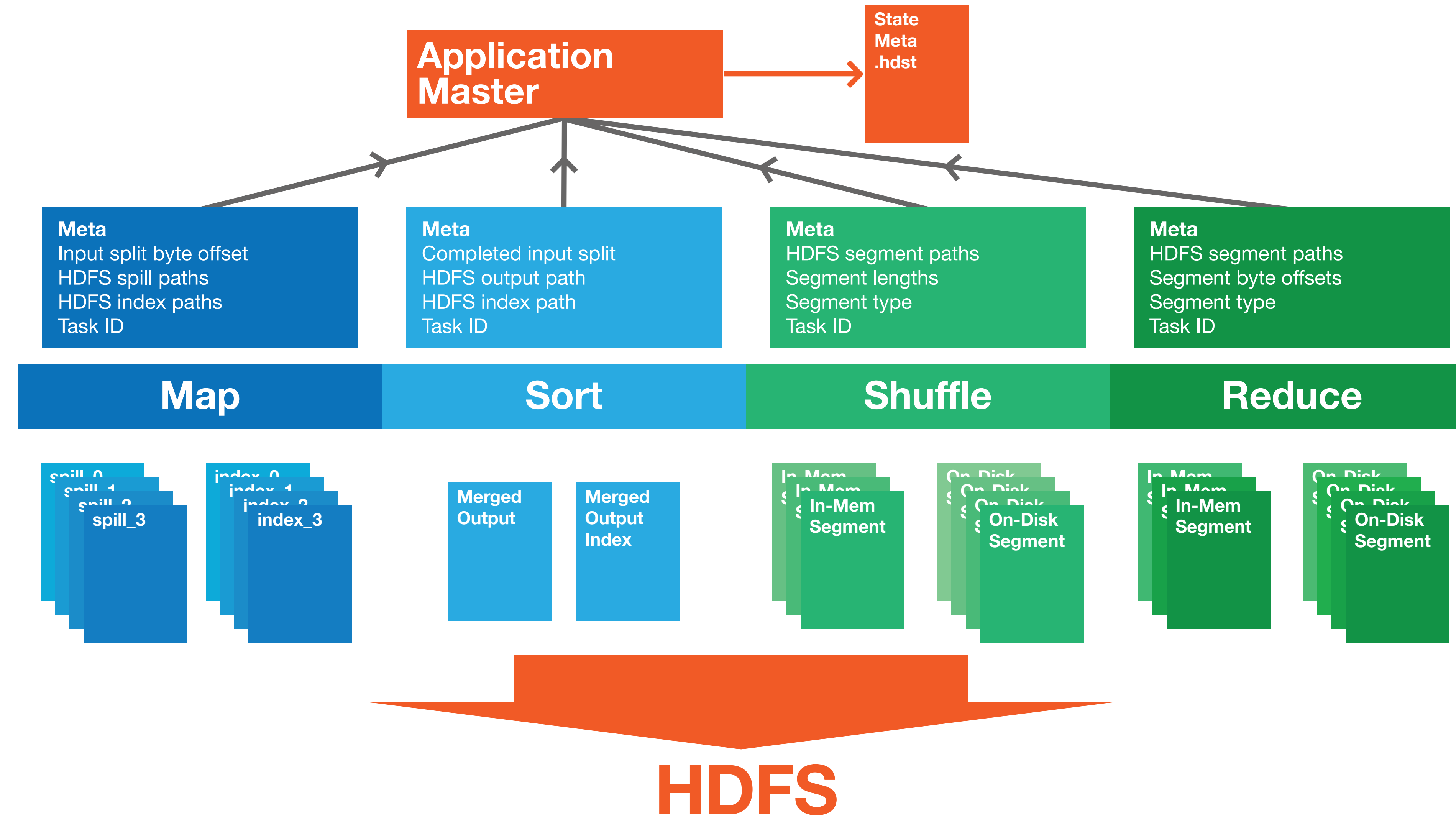
Department of Mathematics & Computer Science, University of Puget Sound, Tacoma WA, 98416

## motivation

- To allow higher-priority Hadoop jobs to take precedent over running jobs
- Ability to pause a job, free up resources, and restart at any time
- Ability to migrate a job to another cluster
- Exploit volatile energy prices between multiple clusters to save money by shifting load

## design

- Assuming relevant data exists on both clusters
- **Pausing creates a state file in HDFS**
  - Task-level state is calculated independently for each task
  - Each task stores its state to HDFS
  - Each task reports local state metadata to application master
  - Application master saves job-level state metadata to HDFS
- **Restarting a job reads this state file**
  - FileInputFormat generates new splits based on block locations in secondary cluster
  - Application master parses global state metadata and creates new tasks with task-level state metadata
  - Restarting tasks retrieve task-level state files and rebuild data structures



## references

A. Berl, E. Gelenbe, M. D. Girolamo, G. Giuliani, H. D. Meer, M. Q. Dang, and K. Pentikousis, "Energy-efficient cloud computing," *The Computer Journal*, vol. 53, no. 7, pp. 1045–1051, 2009.

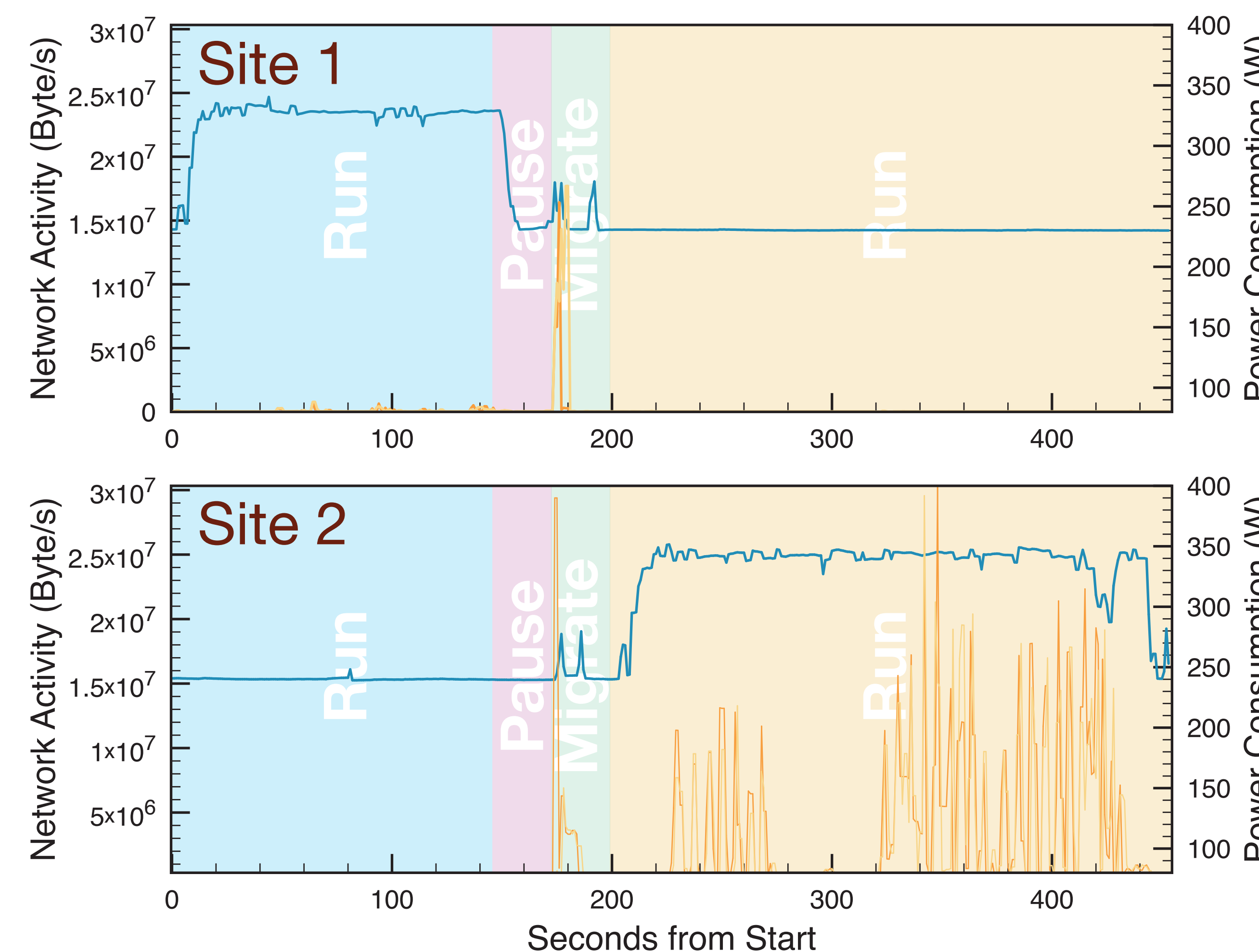
A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future Generation Computer Systems*, May 2011.

S. Akoush, R. Sohan, A. Rice, A. W. Moore, and A. Hopper, "Free lunch: exploiting renewable energy for computing," in *Proceedings of the 13th USENIX Workshop on Hot Topics in Operating Systems (HotOS 2011)*. USENIX, 2011.

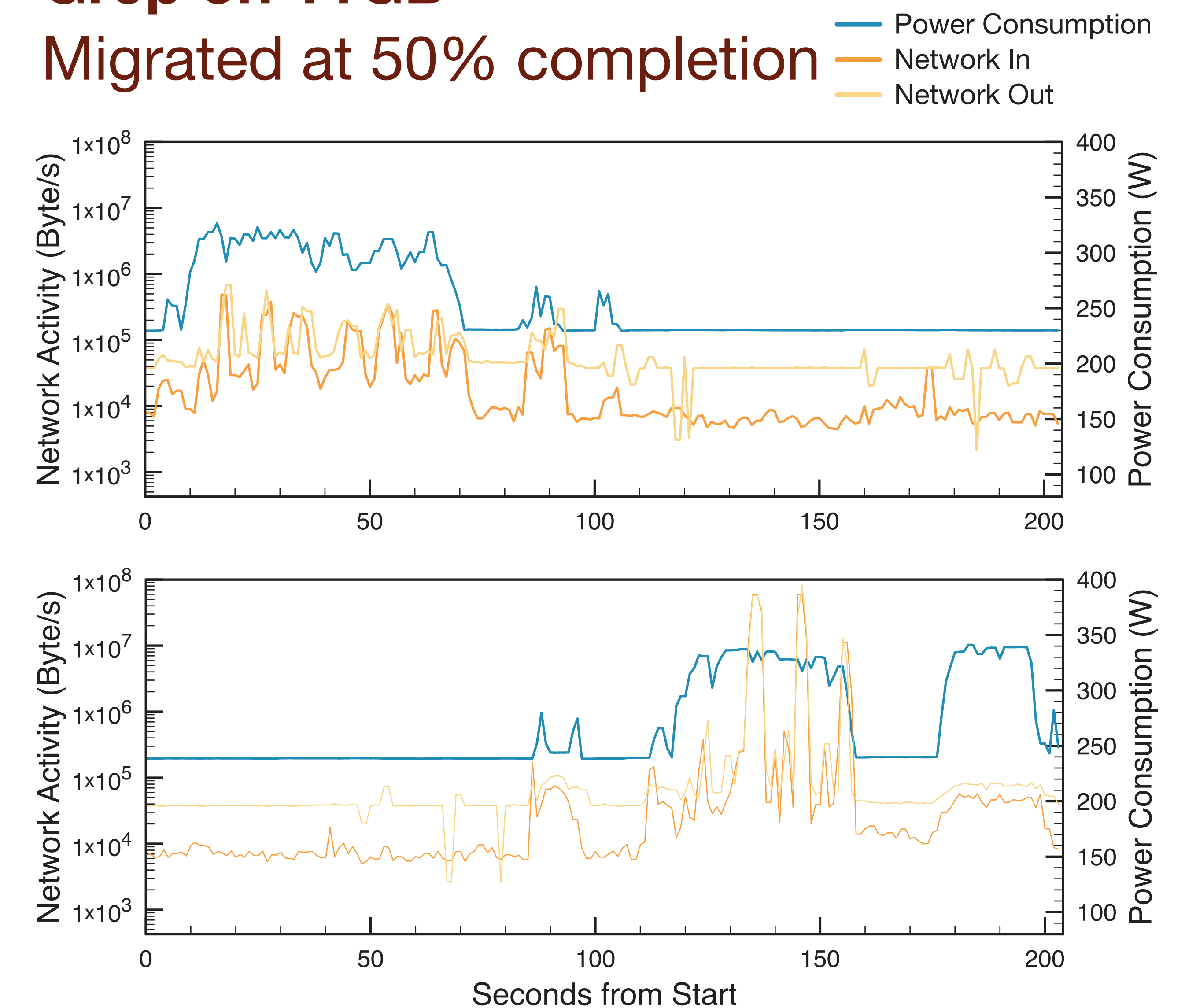
Y. Li, D. Chiu, C. Liu, L. T. X. Phan, T. Gill, S. Aggarwal, Z. Zhang, B. T. Loo, D. Maier, and B. McManus, "Towards dynamic pricing-based collaborative optimizations for green data centers," in *Workshops Proceedings of the 29th IEEE International Conference on Data Engineering, ICDE 2013, Brisbane, Australia, April 8-12, 2013, 2013*, pp. 272–278.

## preliminary results

### Wordcount on 11GB Migrated at 25% completion.



### Grep on 11GB Migrated at 50% completion



**acknowledgments** ■ This research was funded in part by the McCormick Research Grant, University of Puget Sound.