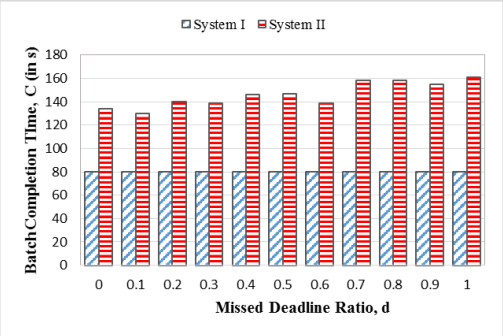

Figure 4a. Impact of Missed Deadline Ratio on Batch Completion Time for Small Workload

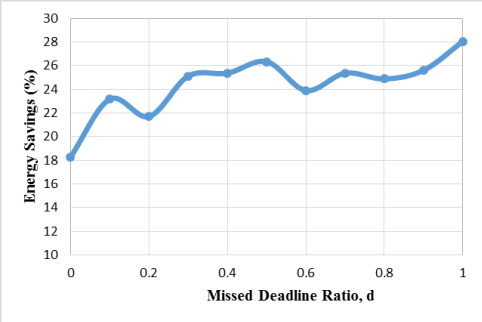

Figure 4b. Impact of Missed Deadline Ratio on Energy Savings for Small Workload

6. SUMMARY AND CONCLUSIONS

This paper introduces an energy aware resource management technique for batch workloads of MapReduce jobs subject to SLAs which include earliest start times, task execution times, and soft deadlines specified by the user. The technique makes use of available slack time in the execution window of jobs and applies a DVFS-based processor frequency reduction to the execution of some tasks to reduce energy consumption *without violating SLAs*. Preliminary performance analysis demonstrates energy savings between 16% and 45% for varying values of missed deadline ratio, laxity factor, and workload size using the DVFS-based approach compared to previous resource management approaches which do not consider energy. Smaller, but still substantial, energy savings were observed for workloads with fewer total tasks. The energy savings were accompanied by small – moderate increases in batch completion times. Plans for future research include:

- Extensively evaluating performance impact of varying workload size and processor model including large workloads from [9] and alternative processor models presented in [11].
- Extending the DVFS-based approach to handle an open stream of job arrivals subject to SLAs.
- Incorporating MapReduce task execution time prediction models to evaluate the impact of and devise techniques to handle error in execution times estimates of MapReduce tasks.
- Enhancing the DVFS-based approach to consider data locality and data transfer times for performance measurements of an implementation on a real-world Hadoop Custer.

7. REFERENCES

[1] Cardosa, M., Singh, A., Pucha, H., Chandra, A., "Exploiting Spatio-temporal Tradeoffs for Energy- Aware MapReduce in the Cloud," *Cloud Computing (CLOUD), 2011 IEEE International Conference on* , vol., no., pp.251,258.

[2] Dean, J. and Ghemawat, S. 2004. MapReduce: Simplified data processing on large clusters. International Symposium on Operating System Design and Implementation (December 2004). 137–150.

[3] Íñigo Goiri, Josep Ll. Berral, J. Oriol Fitó, Ferran Julií, Ramon Nou, Jordi Guitart, Ricard Gavaldí, and Jordi Torres. 2012. Energy-efficient and multifaceted resource management for profit-driven virtualized data centers. Future Gener. Comput. Syst. 28, 5 (May 2012), 718-731

[4] Hamilton, J. "Cooperative expendable micro-slice servers(cems): low cost, low power servers for internet-scale services," in *Proc. of the Conf. on Innovative Data Systems Research,* 2009.

[5] Hewlett-Packard Corporation, Intel Corporation, Microsoft Corporation, Phoenix Technologies Ltd., and Toshiba Corporation, Advanced Configuration and Power Interface Specification, Revision 5.0a, December 6, 2011.

[6] IBM. IBM ILOG CPLEX Optimization Studio. Available: http://www-03.ibm.com/software/products/us/en/ibmilogcpleoptistud

[7] Kyong Hoon Kim, Buyya, R., Jong Kim, "Power Aware Scheduling of Bag-of-Tasks Applications with Deadline Constraints on DVS-enabled Clusters," in Cluster Computing and the Grid, 2007. CCGRID 2007. vol., no., pp.541-548, 14-17 May 2007

[8] Koomey, J. 2011. Growth in data center electricity use 2005 to 2010. Oakland, CA: Analytics Press. August, vol . 1.

[9] Lim, N, Majumdar, S, and Ashwood-Smith, P. 2014. "Engineering Resource Management Middleware for Optimizing the Performance of Clouds Processing MapReduce Jobs with Deadlines", in Proc. 5Th ACM/SPEC ICPE Dublin, March 2014.

[10] Rossi, F, van Beek, P, and Walsh, T. 2004. "Chapter 4: Constraint Programming," in Handbook of Knowledge Representation, San Diego, CA: Elsevier Science. 181-211.

[11] Tan, L and Chen, Z. 2015. Slow down or halt: Saving the optimal energy for scalable HPC systems. *in Proc. ICPE, 2015*, pp. 241–244.

[12] Verma, A.; Cherkasova, L.; Kumar, V.S.; Campbell, R.H., "Deadline-based workload management for MapReduce environments: Pieces of the performance puzzle," *Network Operations and Management Symposium (NOMS), 2012 IEEE* , vol., no., pp.900,905, 16-20 April 2012

[13] Wang, L, von Laszewski, G, Dayal, J, and Wang, F. 2010. "Towards energy aware scheduling for precedence constrained parallel tasks in a cluster with dvfs," in Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, ser. CCGRID '10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 368–377.

[14] Wirtz, T and Ge, R. 2011. Improving MapReduce energy efficiency for computation intensive workloads. *Green Computing Conference and Workshops (IGCC), 2011 International*, vol., no,pp.1,8, 25-28.

[15] Yanfei Li, Ying Wang, Bo Yin, Lu Guan, "An energy efficient resource management method in virtualized cloud environment," *Network Operations and Management Symposium (APNOMS), 2012 14th Asia-Pacific* , vol., no., pp.1,8, 25-27 Sept. 2012