by 25.5%. As a result, 6% more DRAM energy is consumed. With CR4, however, DRAM power and energy consumptions are reduced by 7.1% and 4.7%, respectively, on average but execution time is increased by 2.9% on average. Performance degradation comes from increased rank contention in active state ranks because write traffic to active ranks is increased. With W, execution time, DRAM power and energy consumptions are reduced by 1.7%, 9.8% and 11.2%, respectively, on average.

For the result of address and data bus energy consumption, with the traffic skew scheme, address and data bus energy consumptions are increased by 2.4% and 2.8%, respectively, on average because of increased L2 cache miss rate. CR4 reduces address and data bus energy consumptions by 1.3% and 3.4%, respectively, on average. These results are not directly proportional to reduced writes because bus energy consumption depends on previously transferred bus data. W reduces address and data bus energy consumptions by 6.6% and 1.9%, respectively, on average. W batch-writes write requests to the same rank, bank, row, and column so that address bus energy of bank, row, or column address is reduced.

From the above results, we observe that CR4 achieves DRAM power and energy gain from reduced write requests and state transitions. However, increased execution time (2.9%) of CR4 is much less than increased execution time (25.5%) of the traffic skew scheme. W achieves DRAM power and energy gain from reduced rank state transitions, which comes from the rank-state aware batch-writing.

5. RELATED WORK

There are many approaches proposed to effectively use row buffers. DRAM-aware last-level cache write-back is proposed in [4]. Its key idea is to aggressively send out last-level cache blocks that are expected to hit in DRAM row buffers. The Virtual Write Queue is proposed in [5]. Its key idea is to send out last-level cache blocks which can be written to the same DRAM row buffers as those of the write requests already in progress. These two schemes increase write requests to DRAM because dirty cache blocks, which are mapped to open row buffers, are written to DRAM by cleaning. Adaptive power and thermal management scheme is proposed in [14]. Reducing both row activation energy and DRAM peak temperature are the goals of these schemes by buffering write requests. Buffered write requests are executed when the corresponding row buffer is open. Prefetching-based memory traffic clustering to lengthen DRAM idle times for energy reduction is proposed by [16]. Adaptive history-based memory schedulers are proposed in [12]. They consider the history of recently scheduled operations when the memory controller selects next operation. Memory controller policy for DRAM power management is proposed in [13]. They demonstrate that an idle DRAM chip immediately switching to the low-power state is better than sophisticated policies.

6. CONCLUSIONS

We proposed DRAM power-aware rank scheduling schemes which can be implemented in the last-level cache and the memory controller. We demonstrated that our two schemes reduce DRAM power and energy consumptions and that their overheads are negligible. Our contributions are as follows. First, we reduce write requests to DRAM with negligible overhead. Our scheme utilizing the last-level cache reduces write requests, and DRAM power and energy consumptions by 6.8%, 7.1% and 4.7%, respectively, on average. Second, we increase rank idleness and decrease state transitions. Write requests are delayed in order to maximize rank idleness. Our scheme utilizing the memory controller reduces state transitions, DRAM power and energy consumptions by 21.2%, 9.8% and 11.2%, respectively, on average, with no performance degradation.

7. ACKNOWLEDGMENTS

This work was performed when Sukki Kim was with Dept. of Computer Science at KAIST. This research was supported by the National Research Foundation of Korea (NRF) grants funded by the Ministry of Education, Science and Technology (2011-0005378, 2012-0000980).

8. REFERENCES

[1] C. Lefurgy, K. Rajamani, F. Rawson, W. Felter, M. Kistler, and T. W. Keller, “Energy Management for Commercial Servers”, *IEEE Computer*, vol. 36, pp. 39-48, December 2003.

[2] H. Huang, K. G. Shin, C. Lefurgy, and T. Keller, "Improving Energy Efficiency by Making DRAM Less Randomly Accessed", In *Proc. of the International Symposium on Low Power Electronics and Design*, 2005.

[3] A. M. Amin and Z. A. Chishti, "Rank-Aware Cache Replacement and Write Buffering to Improve DRAM Energy Efficiency", In *Proc. of the International Symposium on Low Power Electronics and Design*, 2010.

[4] C. Lee, V. Narasiman, E, Ebrahimi, O. Mutlu, and Y. N. Patt, “DRAM-Aware Last-Level Cache Writeback: Reducing Write-Caused Interference in Memory Systems”, *HPS Technical Report*, TR-HPS-2010-002, April 2010.

[5] J. Stuecheli, D. Kaseridis, D. Daly, H. C. Hunter, and L. K. John, "The Virtual Write Queue: Coordinating DRAM and Last-Level Cache Policies", In *Proc. of the International Symposium on Computer Architecture*, 2010.

[6] E. Larson and S. Chatterjee, “MASE: A Novel Infrastructure for Detailed Microarchitectural Modeling”, In *Proc. of the International Symposium on Performance Analysis of Systems and Software*, 2001.

[7] P. Rosenfeld, E. Cooper-Balis, and B. Jacob, "DRAMSim2: A Cycle Accurate Memory System Simulator", *Computer Architecture Letters*, vol. 10, no. 1, pp. 16-19, 2011.

[8] Standard Performance Evaluation Corporation, “SPEC CPU 2000 Benchmark Suite”, http://www.specbench.org/osg/cpu2000/

[9] Z. Herczeg, D. Schmidt, A. Kiss, N. Wehn, and T. Gyimothy, "Energy Simulation of Embedded XScale Systems with XEEMU", *Jornal of Embedded Computing*, vol. 3, issue 3, 2009.

[10] JEDEC, “JEDEC Standard DDR3 SDRAM Specification”, http://www.jedec.org/sites/default/files/docs/JESD79-3E.pdf

[11] Micron Inc. “1Gb: x4, x8, x16 DDR3 SDRAM”, http://www.micron.com/products/dram/ddr3-sdram

[12] I. Hur and C. Lin, "Adaptive History-Based Memory Schedulers", In *Proc. of the International Symposium on Microarchitecture*, 2004.

[13] X. Fan, C. Ellis, and A. Lebeck, "Memory Controller Policies for DRAM Power Management", In *Proc. of the International Symposium on Low Power Electronics and Design*, 2001.

[14] S. Liu, S. O. Memik, Y. Zhang, and G. Memik, "A Power and Temperature Aware DRAM Architecture", In *Proc. of the Design Automation Conference*, 2008.

[15] S. Rixner, W. J. Dally, U. J. Kapasi, P. Mattson, and J. D. Owens, “Memory Access Scheduling”, In *Proc. of the International Symposium on Computer Architecture*, 2000.

[16] Y. Lee and S. Kim, “DRAM Energy Reduction by Prefetching-Based Memory Traffic Clustering”, In *Proc. of the Great Lakes Symposium on VLSI*, 2011.