

# Correlating Hardware Performance Events to CPU and DRAM Power Consumption

Michael Giardino and Bonnie Ferri

School of Electrical and

Computer Engineering

Georgia Institute of Technology

Atlanta, Georgia 30332-0250

Email: {giardino, bonnie.ferri}@ece.gatech.edu

## I. INTRODUCTION

There are numerous ways to control power usage of modern systems. Dynamic voltage and frequency scaling (DVFS) and C- and P-states are commonly available in hardware, and operating systems have fine-grained control of these states. Unfortunately, most DVFS systems simply consider CPU utilization when determining which P-state to enter (e.g. Linux's ondemand governor). In addition, even when more intelligent systems have been developed, a limitation in previous research has been how to measure power consumption in real time. Running average power limit (RAPL), introduced by Intel in their Sandy Bridge line of processors, allows researchers and system designer to obtain detailed estimates of energy consumption by the core, uncore and DRAM. Previous research has relied heavily on models to validate power-performance trade-offs. We now have the ability to measure cumulative energy usage in joules calculate average power. In this paper we use Linux's perf\_events subsystem to measure a number of commonly cited performance metrics while running the SPEC CPU2006 benchmarks in addition to calculating average power using RAPL registers. This paper shows initial promising results correlating DRAM memory consumption to stall-cycle ratio. Additional data for other performance metrics (IPC, LLC misses, etc.) will be presented in a poster.

In 2000, Bellosa demonstrated a link between performance events and power consumption [1]. Determining which metrics are most useful for performance and power modeling is an open question.

Lee et al. estimate performance at different frequencies using a regression model that includes measured CPI and the ratio of memory bus instructions to total retired instructions, similar to our Memory Ratio [2]. Rountree et al. explore a number of different performance counters and suggest that *leading loads* (i.e. the first load instruction that misses the LLC) are an effective potential PMC not yet implemented. [3]

Integrating dynamic power data to a system is a challenge. Noureddine et al. wrote a survey of computer energy measurement methods but these predominantly consist of modeling and estimation [4]. These still depend on power modeling/estimation and tend to focus on using a single node

of a cluster as representative.

David et al. introduce the RAPL algorithm for enforcing memory power limits [5]. Rountree et al. explore the actual hardware-enforced power bounds of RAPL and measure the power variance between nodes of large multi-node clusters [6]. Subraminium uses RAPL to evaluate power and performance of SPECweb benchmarks using power limits [7].

## II. TESTING METHODOLOGY

Our experiments were performed on a four-core Intel i7-4770 Haswell CPU. Since the `rdmsr` instruction needed to measure the energy must be run from ring 0, a daemon was written that accumulates total energy and checks the counter for overflow.

The SPEC CPU2006 [8] benchmarks were chosen due to their widespread use in the literature for power modeling.

We ran the 29 benchmarks at 6 different frequencies (1GHz, 1.7GHz, 2.1GHz, 2.7GHz, 3.1GHz, and using the `on_demand` governor) five times each to collect data. Four independent copies of each benchmark were run in parallel to fully load the system and each one was pinned to a physical core.

## III. RESULTS

To determine whether existing performance metrics are useful for estimating overall power usage, we conducted a number of experiments measuring both core and DRAM energy usage. We defined relative execution time as  $RET_f = t_f/t_{max}$  where  $RET_f$  is the relative execution time at frequency  $f$ ,  $t_f$  is execution time at frequency  $f$  and  $t_{max}$  is the maximum execution time (which in our experiments always occurred at the lowest frequency).

Among the metrics we tested, the most useful metric for estimating power usage is stall cycles. We used stall cycles per total cycles to get a ratio between zero and one. Core power and DRAM power are shown as Figures 1 and 2 respectively. As the stall cycle ratio decreases, the core power increases nearly monotonically. The DRAM power consumption has much more noise, though the trend certainly suggests that as the stall cycle ratio increases, the memory power consumption increases. In addition, Figure 3 shows the relationship between stall cycle ratio and relative execution time. It very clearly

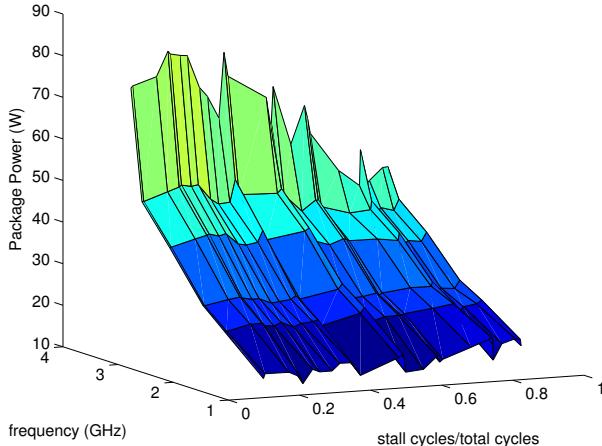


Fig. 1. Stall Cycles Per Cycle vs. Core Power

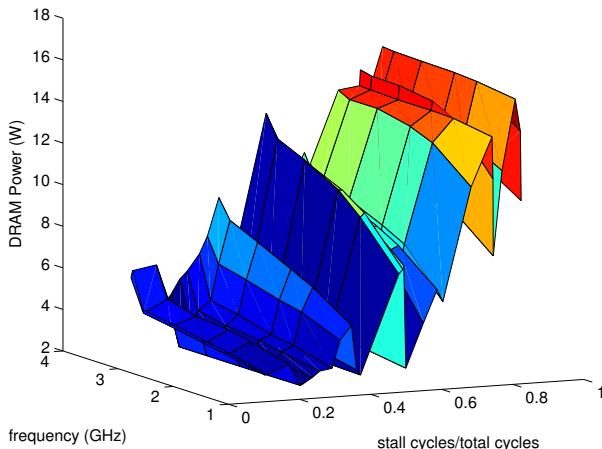


Fig. 2. Stall Cycles per Cycle vs. DRAM Power

shows that with a stall cycle ratio greater than 0.5, the potential speedup from increasing core frequency is greatly reduced. If fewer than half of total cycles are stall cycles, there is at least a 50% speedup potential from the lowest frequency to the highest frequency.

As expected, core and DRAM power consumption were for the most part inversely proportional. We found that the less commonly used performance event stall cycles gave the best estimate of power consumption: the higher the stall ratio, the more likely memory power is going to take a greater portion of your power budget and thus it may be worthwhile to be more aggressive in reducing the frequency of the core. Conversely, if the stall cycle ratio is low, memory power is very low, suggesting aggressive DRAM power-down schemes may be employed.

Similar data were collected for CPI, memory instruction ratio, and LLC misses but was omitted due to space constraints.

We are exploring compound/derived metrics as well as a

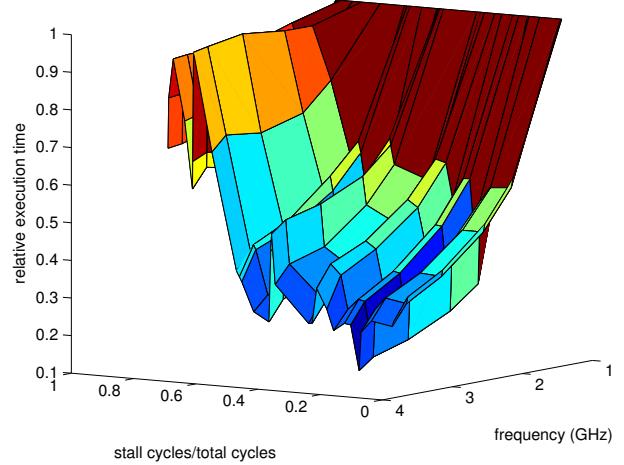


Fig. 3. Stall Cycles per Cycle vs. Relative Execution Timer

deeper statistical analysis of the data. We are also developing models to allow systems without RAPL-type capabilities to accurately estimate power consumption based upon metrics available to the system. In addition, we are developing on a kernel module for measuring power in real-time in order to obtain a time-domain picture of how the performance events relate to memory and core power consumption in order to better control DVFS and sleep states.

## REFERENCES

- [1] F. Bellosa, “The benefits of event: Driven energy accounting in power-sensitive systems,” in *Proceedings of the 9th Workshop on ACM SIGOPS European Workshop: Beyond the PC: New Challenges for the Operating System*, ser. EW ’9. New York, NY, USA: ACM, 2000, pp. 37–42. [Online]. Available: <http://doi.acm.org/10.1145/566726.566736>
- [2] S.-J. Lee, H.-K. Lee, and P.-C. Yew, “Runtime performance projection model for dynamic power management,” in *Proceedings of the 12th Asia-Pacific Conference on Advances in Computer Systems Architecture*, ser. ACSAC’07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 186–197. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2392163.2392182>
- [3] B. Rountree, D. Lowenthal, M. Schulz, and B. De Supinski, “Practical performance prediction under dynamic voltage frequency scaling,” in *Green Computing Conference and Workshops (IGCC), 2011 International*, July 2011, pp. 1–8.
- [4] A. Noureddine, R. Rouvoy, and L. Seinturier, “A review of energy measurement approaches,” *SIGOPS Oper. Syst. Rev.*, vol. 47, no. 3, pp. 42–49, Nov. 2013. [Online]. Available: <http://doi.acm.org/10.1145/2553070.2553077>
- [5] H. David, E. Gorbato, U. R. Hanebutte, R. Khanna, and C. Le, “Rapl: Memory power estimation and capping,” in *Proceedings of the 16th ACM/IEEE International Symposium on Low Power Electronics and Design*, ser. ISLPED ’10. New York, NY, USA: ACM, 2010, pp. 189–194. [Online]. Available: <http://doi.acm.org/10.1145/1840845.1840883>
- [6] B. Rountree, D. Ahn, B. De Supinski, D. Lowenthal, and M. Schulz, “Beyond dvfs: A first look at performance under a hardware-enforced power bound,” in *Parallel and Distributed Processing Symposium Workshops PhD Forum (IPDPSW), 2012 IEEE 26th International*, May 2012, pp. 947–953.
- [7] B. Subramaniam, “Metrics, models and methodologies for energy-proportional computing,” in *Cluster, Cloud and Grid Computing (CC-Grid), 2014 14th IEEE/ACM International Symposium on*, May 2014, pp. 575–578.
- [8] J. L. Henning, “Spec cpu2006 benchmark descriptions,” *SIGARCH Comput. Archit. News*, vol. 34, no. 4, pp. 1–17, Sep. 2006. [Online]. Available: <http://doi.acm.org/prx.library.gatech.edu/10.1145/1186736.1186737>