

Effects of Active Cooling on Workload Management in High Performance Processors

Won Ho Park and C. K. Ken Yang

Department of Electrical Engineering, UCLA, Los Angeles, U.S.A.
wonhopark80@gmail.com, yang@ee.ucla.edu

Keywords: Power-aware Systems, Workload Scheduling, Server Provisioning, Electronic Cooling, Data Centers, Thermal Management.

Abstract: This paper presents an energy-efficient workload scheduling methodology for multi-core multi-processor systems under actively cooled environment that improves overall system power performance with minimal response time degradation. Using a highly efficient miniature-scale refrigeration system, we show that active-cooling by refrigeration on a per-server basis not only leads to substantial power-performance improvement, but also improves the overall system performance without increasing the overall system power including the cost of cooling. Based on the measured results, we present a model that captures different relations and parameters of multi-core processor and the refrigeration system. This model is extended to illustrate the potential of power optimization of multi-core multi-processor systems and to investigate different methodologies of workload scheduling under the actively cooled environment to maximize power efficiency while minimizing response time. We propose an energy-efficient workload scheduling methodology that results in total consumption comparable to the *spatial subsetting* scheme but with faster response time under the actively cooled environment. The actively cooled system results in $\geq 29\%$ of power reduction over the non-refrigerated design across the entire range of utilization levels. The proposed methodology is further combined with the G/G/m-model to investigate the trade-off between the total power and target SLA requirements.

1 INTRODUCTION

Increase in energy consumption due to the tremendous growth in the number and size of data centers presents a whole new set of challenges in maintaining energy-efficient infrastructure. While data centers' energy consumption had accounted for 2% of the total energy budget of the USA in 2007, it is expected to reach 4% by the year 2011. This number is equivalent to \$7.4 billion per year on electric power where this number has changed by 60% since 2006 (U.S. EPA, 2007).

Worldwide trend of energy consumption in data centers tracks the US trend (Rajamani, 2008). Fig. 1 shows the number of data center installations, worldwide new server spending, and electric power and cooling costs. Despite the steady increase of installed base of data centers over the last decade, new server spending has stayed relatively constant due to the decrease in electronic costs. As the data center infrastructure becomes denser, power density has been increasing by approximately 15% annually

(Humphreys, 2006), hence increasing electricity consumption for operating servers and cooling. It is likely that IT operating cost will soon outweigh the initial capital investment.

Detailed energy breakdown of different types of data centers can be found in (Rajamani, 2008), (Tschudi, 2003), (Lawrence Berkeley National Labs, 2007), (Patterson, 2008). A data center can consist of hundreds or even thousands of server racks where each rack can draw more than 20kW of power. Relative percentage of various contributors to energy usage varies considerably among data centers, but up to 90% of the total energy is attributed to the energy dissipated by the computer load and the energy required by the Computer Room Air Conditioning unit (CRAC). Additionally, there is a strong relation between the energy consumed by the computer load and the CRAC units since any reduction in electronic heat can be compounded in the cooling system. For example, CRAC energy efficiency of data centers can increase by 1% per degree Celsius.

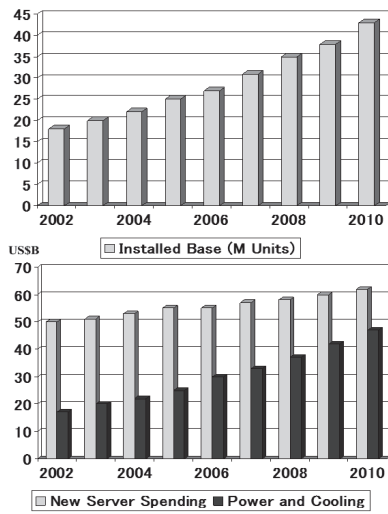


Figure 1: (a) Number of worldwide installed bases of data center. (b) Worldwide spending on new servers and operation cost. Adopted from (Rajamani, 2008).

Reducing the overall energy usage is an area of interest across multiple disciplines. The focus of most efforts on energy/power saving in server systems is on processor elements (Jing, 2011), (Chaparro, 2007), (Ma, 2003), (Tschanz, 2003), (Brooks, 2000), (Sato, 2007), (Ghosh, 2011), (Rabaey, 2003). Approaches for power saving of processors often adopt both the software-based energy-aware workload scheduling (Jing, 2011), (Lin, 2011), (Luo, 2013) and hardware-based circuit and architectural power management techniques to effectively optimize energy usage. A typical software-based workload scheduling algorithm controls energy by distributing workloads to processors in a way to reduce both the electric and cooling costs. The basis of these approaches relies on powering-off servers that are not utilized by concentrating the workload on a subset of the servers. This method is known as *spatial subsetting*, and has been shown to successfully tackle the issue of idle server power consumption (Jing, 2011), (Pinheiro, 2001), (Chase, 2001).

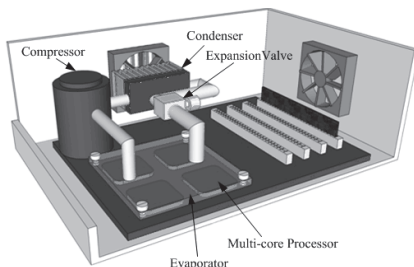


Figure 2: Configuration of a multi-processor computing server unit with a refrigerated-cooling.

Moreover, energy savings from the off-power servers is compounded in the cooling systems that consume power to remove the heat dissipated in the servers. While this approach significantly reduces idle power, it raises a concern of degraded response time in computing systems, due to the power-latency trade off.

To address the problem of degraded response time in *spatial subsetting*, one solution is to employ an *over-provisioning* scheme (Chen, 2005), (Ahmad, 2010). The over-provisioning algorithm can be considered as a power and response time optimization problem. By predicting how many servers are required to service the requested workload, the workload management software assigns a subset of processors to remain at idle state to absorb sudden increases in the load. Determining the number of server to be held at idle state often relies on a good model that successfully plans capacity depending on the upcoming workloads. For instance, G/G/m-models from queueing theory have been used to obtain useful measures like average execution velocity and average wait time to support capacity and workload planning of multi-processor systems in order to satisfy target SLA requirements (Chen, 2005), (Ahmad, 2010), (Müller-Colstermann, 2007).

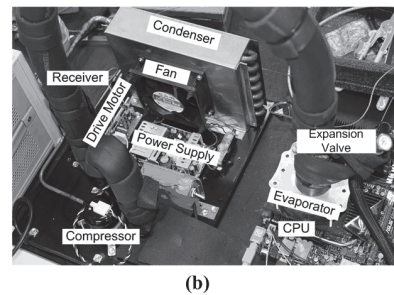
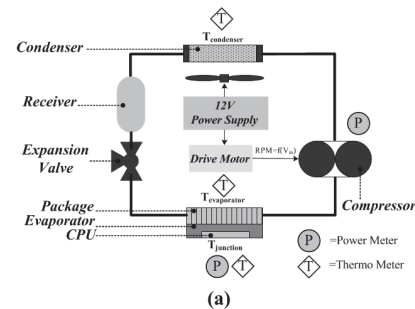


Figure 3: (a) Layout and (b) photograph of the refrigeration system for electronic cooling.

Often the software utilizes special hardware supports (Jing, 2011), such as dynamic voltage frequency scaling (McGowen, 2006), (Burd, 2000),

(Nowka, 2002), or thread migration (Zhang, 2005), stopping a processor through power gating (Tschanz, 2003), (Zhang, 2005), (Henzler, 2005), or body biasing (Tschanz, 2003). However, these techniques not only induces area penalty but also require some transition time in and out of the low-power state and imply performance degradation.

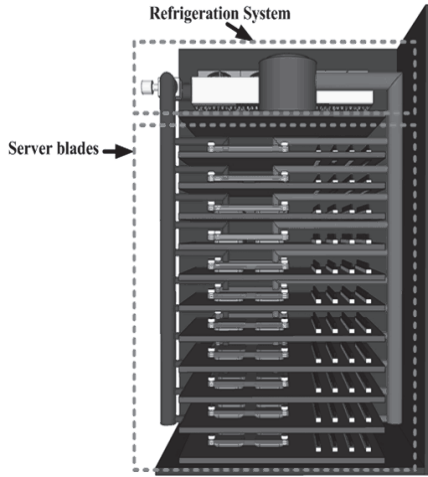


Figure 4: Configuration of a multi-processor computing server unit with a refrigerated-cooling.

Along with these techniques, actively cooling the processors using refrigeration has attracted recent interest as a practical option to ease the power problems in high performance computing units (Copeland 2005), (Mahajan, 2006), (Nnanna, 2006), (Chu, 2004), (Trutassanawin, 2006). Operating CMOS circuitry at sub-ambient temperatures for higher performance has been shown over the past few decades (Carson, 1989), (Aller, 2000). While the speed improvement can be traded for lower power dissipation of the electronics, the cost of cooling can limit the overall system power performance. Recent work (Park, 2010), (Park, 2010), has shown that active cooling not only can lead to overall power improvement that includes the cost of cooling power without performance degradation; the results show that the amount of power savings is roughly proportional to the ratio of leakage power to total power due to the exponential sensitivity of leakage power to temperature, irrespective of type of workload. For instance, cooling a processor that dissipates 175.4W of power with 30% electronic leakage power resulted in a total system power consumption of 133W. This performance is 25% better than the non-cooled reference design (Park, 2010). Focus of this paper is to explore the effectiveness of workload scheduling to improve power efficiency of multi-core multi-

processor systems in an actively cooled environment using a highly efficient refrigeration system. Results presented in this paper suggest that there exists a methodology under actively cooled environment that optimizes power efficiency while minimizing response time in and out of the low-power state. Furthermore, we combine our proposed methodology with the G/G/m-models to reduce both total power and response time degradation while meeting target SLA requirements.

2 MULTI-CORE PROCESSOR UNDER THE ACTIVELY-COOLED ENVIRONMENT

A miniature-scale refrigeration system for electronic cooling that is capable of operating at a reduced temperature with high efficiency has been developed and experimentally tested in (Park, 2010). The compressor used in our miniature refrigeration system has cooling capacity in the several hundred-watt ranges, indicating that this refrigeration system can potentially be configured to simultaneously cool multi-processor servers. We envision a possible configuration of the HPC server unit as illustrated in Fig. 2.

A layout and photograph of the refrigeration system for electronic cooling is shown in Fig. 3. A configuration of the refrigeration system charged with R-134a refrigerant consists of a compressor, condenser, an expansion valve, a cold plate, evaporator, and a cooling fan. A 12V power supply provided the required power. Additionally, a motor drive board is installed to control the compressor speed and modulate the refrigeration capacity at different loads. K-type bead probes are taped to the evaporator and the condenser for temperature measurements. Power meters are used to measure power consumptions of the cooler and the heat source. By controlling the speed of the compressor, we cool the microprocessor at different heat loads and temperatures in order to obtain minimal total system power. Specific chip junction temperature would be the temperature that resulted in the lowest system power. The detailed description of the experimental setup and performance of our miniature refrigeration system for electronic cooling can be found in (Park, 2010). We characterize the power performance of a 4-core processor at different operating conditions using this refrigeration system. It is important to mention that while our analysis

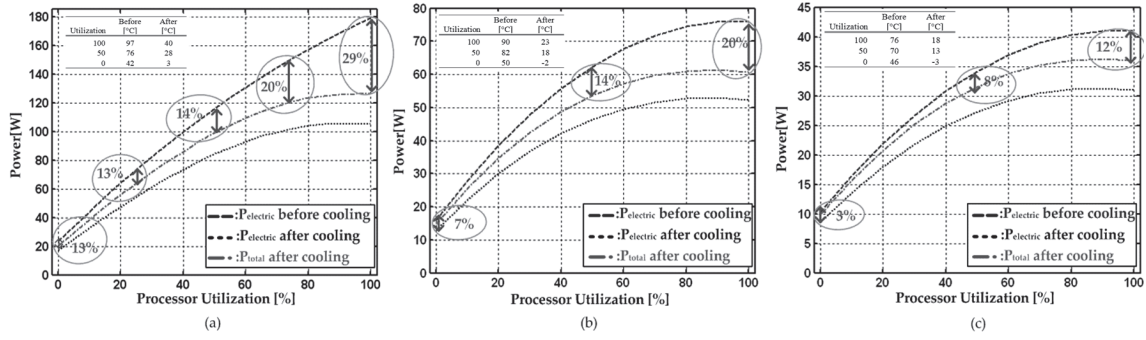


Figure 5: Power consumption before and after cooling across different processor utilization levels when (a) 4, (b) 2, or (c) 1 core out of the 4-core processor is powered up based on the model and measured data. The associated power savings after electronic cooling is also shown.

uses a system that can be enclosed in a server chassis, and vapor compression refrigeration systems can achieve considerably higher efficiency with larger cooling capacity at the expense of larger volume. Such systems can potentially cool entire racks of servers with the coolant distributed with parallel flow through the server blades as shown in Fig. 4. The results discussed in this paper can be directly applied.

The mechanisms for power dissipation of digital CMOS ICs are well understood. The total power dissipation can be estimated by the sum of the active power and leakage power (Rabaey, 2003), (Chandrakasan, 1992).

$$P_{electric} = P_{active} + P_{leakage} \quad (1)$$

$$P_{electric} = \alpha * C_{switched} * f_{clk} * V_{dd}^2 \quad (2)$$

$$P_{leakage} = V_{dd} * I_0 \exp\left(\frac{-V_{th}}{kT_{junction}/q}\right) \quad (3)$$

The active power, P_{active} depends on the activity factor, α , and the amount of power that dissipates charge/discharge capacitive nodes between the supply voltage (V_{dd}) and ground when executing the logic, $C_{switched}f_{clk}V_{dd}^2$. At nano-meter scale technology, the switches that implement the logic results in a leakage current to flow through each logic gate even when the logic is not active. This leakage becomes a significant component of total chip power in modern era processors. The leakage power, $P_{leakage}$, has an exponential relation with the degree that a transistor's ON/OFF threshold, V_{th} , exceeds the thermal voltage, $KT_{junction}/q$. The $P_{leakage}$ equation simplifies the dependence of leakage power by lumping (1) the number and size of logical switching paths in a computational unit, (2) the carrier properties in the transistor, and (3)

dependence of leakage on the logical structure of each logic gate of a digital processor into a single constant I_0 .

For a digital processor, power dissipation and computing performance are closely related. The Equation (4) shows this relationship for the delay of a logic gate. The current is a function of temperature and primarily depends on the carrier mobility. A designer can typically trade-off any improved speed performance by reducing the supply voltage, V_{dd} .

$$Delay \propto RC \propto \frac{V_{dd}}{I_{logic}} \quad (4)$$

Lower temperatures lead to improved performance of electronic devices. Lower power and higher speed results from (1) an increase in carrier mobility and saturation velocity, (2) an exponential reduction in sub-threshold currents from a steeper sub-threshold slope (KT/q), (3) an improved metal conductivity for lower delay, and (4) better threshold voltage control enabling to lower V_{th} .

For coolers and refrigerators, the efficiency is represented in terms of COP defined by

$$COP = \frac{P_{electric}}{P_{cooling}} \quad (5)$$

where $P_{cooling}$ represents the cooling power of the refrigeration system required to lift the total amount of heat ($P_{electric}$) generated by the processor.

Furthermore, the cooling power can be expressed in terms of COP and the COP of the Carnot cycle with Eq. (6) and (7) where T_{evap} is the cold-end temperature of the evaporator, T_{cond} is the temperature at the condenser, and η is the second law of efficiency.

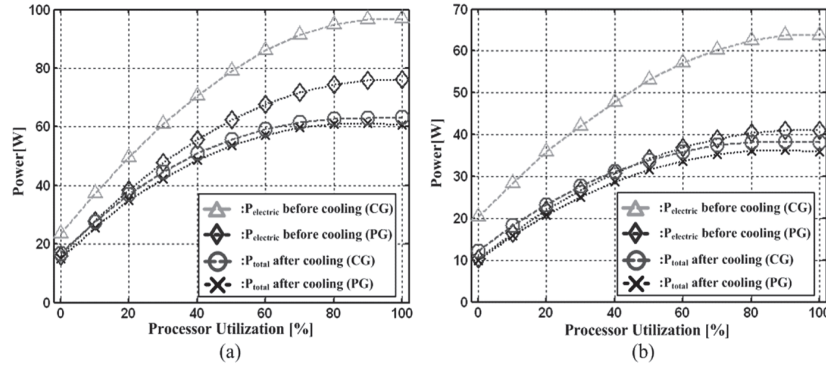


Figure 6: Power consumption across different utilization level before and after cooling using PG and CG as the core stopping techniques for (a) 2-core and (b) 1-core processor.

$$\frac{P_{electric}}{P_{cooling}} = \eta COP_{carnot} = \eta \left(\frac{T_{evap}}{T_{cond} - T_{evap}} \right) \quad (6)$$

$$P_{cooling} = \frac{P_{electric}}{\eta} \left(\frac{T_{cond}}{T_{evap}} - 1 \right) \quad (7)$$

Equating Eq. (1) and (7) results in total system power of

$$P_{total} = P_{electric} + P_{cooling} \quad (8)$$

that includes the cooling power consumption in order to quantify whether the system offers an overall power reduction at different operating temperature. The model serves as a useful tool to evaluate overall system performance including optimal operating temperatures and the amount of total power reduction.

Using this approach, we explore the optimal operating conditions of the system across different processor utilization. In order to experimentally quantify the power consumption of compute-intensive processors, the workload used in all our experiments is Intel's LINPACK, workload, which is CPU bound. Note that our model tracks well with measured data (Park, 2010), (Park, 2010). It is also important to emphasize that our experiment not only uses voltage scaling to trade-off the improved speed performance into a power reduction but also controls refrigeration system to modulate the cooling capacity in order to obtain minimal system power. The speed performance of the processor is kept constant across utilization level. The results are used to build a model of the 4-core processor operating at reduced temperatures and applied to multi-core multi-processors in later sections.

The 4-core processor can be configured such that 1, 2 or 4 cores are active while unused cores are completely turned off to address the problem of idle

power consumption. The refrigeration system is used to cool the microprocessor at different configurations. The amount of total power before and after cooling and the associated power saving across different process utilization levels for different number of cores is shown in Fig. 5. Here, total power before cooling represents forced air cooling that includes the fan power. As can be seen, the result shows that the total power savings of at least 3, 7 and 13 percent can be obtained across the entire range of processor utilization for 1, 2, and 4 cores respectively. The detailed temperature and voltage operating points and the breakdown of the total system power in terms of active, leakage, and cooling components at different utilization with and without active cooling components are shown in (Park, 2010). Effectiveness of cooling is proportional to utilization level. This result suggests that the energy-conscious provisioning would need to concentrate the workload on a minimal active set of cores that run near a maximum utilization level, while other excess cores transition to low-power states to reduce the energy cost. However, using power gating (PG) technique to power on/off cores comes at a price of response time degradation since powering up a core that is completely shut down requires up to 1000 cycles (Kumar, 2003).

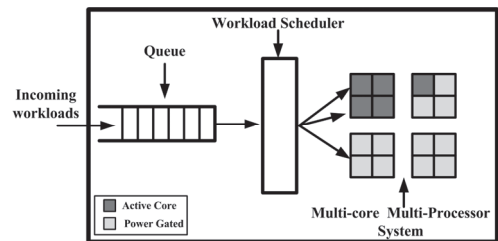


Figure 7: Generic workload scheduling management for multi-core multi-processor computing system.

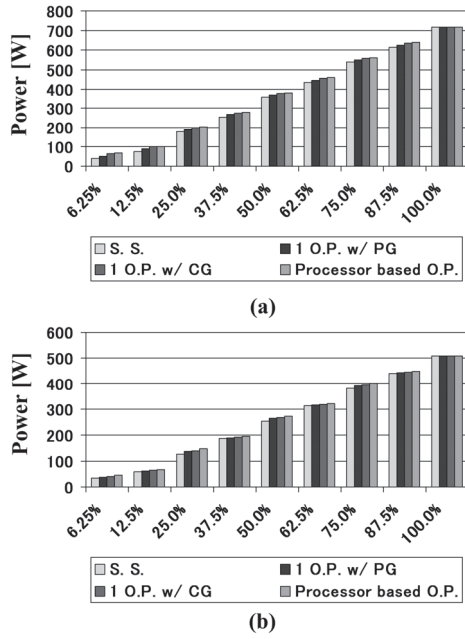


Figure 8: Power at different utilization (a) before and (b) after electronic cooling for different methodologies

On the other hand, a simpler way to stop a core with minimal response time degradation is to clock gate (CG) the core (Tschanz, 2003), (Kurd, 2001). Main advantage of this power saving technique is the state of the processor can be preserved since supply voltage is not cut. This provides a response time which is orders of magnitude faster than waking from power gating or powering up a processor. However, in terms of power consumption, this technique stops dynamic power dissipation, but since power is not entirely cut-off, the core continues dissipating leakage power. Operating CMOS circuitry at reduced temperatures substantially reduces the power since leakage power depends exponentially on temperature. The result that captures the impact of CG at reduced temperatures is shown in Fig. 6.

Before cooling, the CG processor consumes considerably higher power as compared to the PG processor, due to the increase in leakage power. As expected, lowering the temperature of the CG processor exponentially reduces leakage power and results in total power that is comparable to PG processors. At 100% utilization level, power savings from cooling with CG and PG are 36% and 20%, respectively, for a 2-core processor. Results are more significant for a 1-core processor where power savings from cooling with CG and PG are 40% and 12%, respectively. For both cases, CG appears to be a better core stopping technique under the actively

cooled environment. In this way, response time significantly improves at the expense of negligible (~2.5W) power penalty.

The model that captures different relations and parameters of our 4-core processor and the refrigeration system is extended to illustrate the potential of power optimization of multi-core multi-processor systems and investigate different methodologies of workload scheduling under the actively cooled environment.

3 WORKLOAD SCHEDULING METHODOLOGY

With our model derived in Section 2, energy-aware workload scheduling algorithms assign incoming workload to available processors such that power consumption is minimized as constrained by response time requirements.

The server platform we analyze consists of a 4-processor server system with 4-cores per processor under the actively cooled environment. In particular, we are interested in aspects where the effects of electronic cooling change the conventional way of assigning workloads. Detailed results and discussions are presented in this section.

Fig. 7 provides generic management architecture for multi-core multi-processor computing systems where 5 out of 16 cores are utilized. This particular HPC server unit has total of 100% utilization level where each core is responsible for 6.25%. The methodologies we evaluate are the following:

- *Spatial Subsetting (S.S.)*: We assume that unused cores power off by PG. The next core can turn up upon arrival of the workload when the current core is fully occupied.
- *1 Core Over-provision with PG (1 O.P. w/ PG)*: Similar to *spatial subsetting* but one core remains at idle state to absorb sudden peaks in loading.
- *1 Core Over-provision with CG (1 O.P. w/ CG)*: Similar to *1 Core over-provision with PG* but uses CG for the core stopping mechanism.
- *Processor based Over-provision (Processor based O.P.)*: Neither PG nor CG is employed and unused cores remain at idle state.

For all cases, the next processor powers up after all four cores within the active processor are fully utilized to prevent idle power consumption.

For comparison purposes, we show the amount of total power consumption before and after cooling for different types of methodologies across varying

utilization levels in Fig. 8 (a) and (b). Note that the total power after cooling includes the cost of cooling. The impact of cooling on different schemes can be seen through the associated power savings as illustrated in Fig. 9. Several observations can be made based on the results. First, *spatial subsetting* clearly consumes the least amount of power, but the advantage diminishes under the cooled environment. Second, the *processor based over-provision* scheme dissipates the largest amount of power but has no response time degradation. Third, the *1 core over-provision with CG* scheme achieves an excellent compromise that provides the largest amount of power reduction from cooling. Finally, since the next processor powers up after all four cores within the active processor are fully utilized, three power-up transition delays are unavoidable for all cases. They occur from 25% to 31.25%, 50% to 56.25%, and 75% to 81.25%.

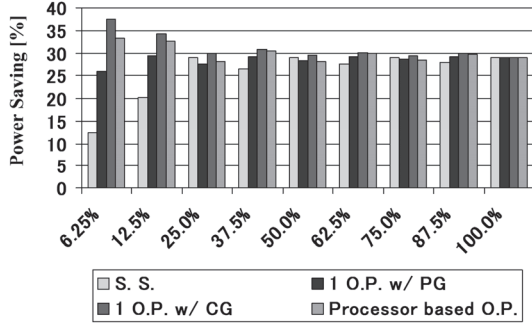


Figure 9: Associated power savings at different utilization level from electronic cooling for different workload assignment methodologies.

Next, we show a new way of assigning workloads under refrigerated cooling and the approach is described in Fig. 10. We demonstrate that the proposed way reduces both the power consumption and the response time requirements at reduced temperature, resulting in power comparable to spatial subsetting but provides a similar response time as 1 core over-provision with CG. Example of the approach is shown in Fig. 10.a; given a workload that requires 4 cores at 100% utilization, the workload scheduling is such that 4 cores are assigned equally to 2 processors. Total power consumption of 196W and 127W is measured, before and after cooling, resulting in a 35% power reduction. On the other hand, the system employing (b) the spatial subsetting scheme and (c) the 1 core over-provision with CG scheme consumes 179W and 199W and 127W and 140W before and after cooling, respectively. The amount of total power saving of the proposed approach is considerably

higher compared to (b) and (c), which has 29% and 30% of power savings.

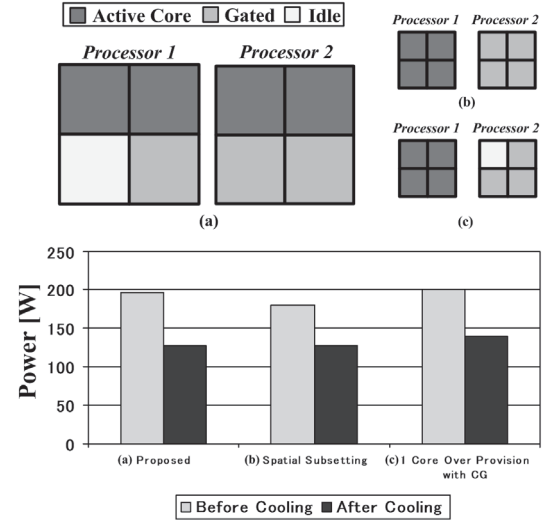


Figure 10: (a) Proposed methodology compared with (b) spatial subsetting and (c) 1 core over-provision with CG.

To be complete, we show the proposed workload scheduling methodology for different utilization levels in Fig. 11. Since power-up events are necessary when a new processor is brought online, three power-up transition delays are unpreventable. These events occur when (B) transitions to (C), (D) transitions to (E), and (F) transitions to (G). In between these transitions and at higher utilizations beyond (G), performance does not degrade with increasing utilization besides the response delay of a few cycles due to CG. Fig. 12 plots the power dissipation and the percentage savings before and after cooling for each of the conditions shown in Fig. 11.

Although conclusions in this section are drawn from a given platform, the intent is not to restrict to a particular platform. The absolute amount of power saving number would be different as different type of systems would have different electronic profile. However, we suggest applying the idea to larger systems where the proposed workload scheduling methodology is applied after cooling. Leveraging the benefits of clock gating at reduced temperatures, our methodology reduces both the power consumption and the response time requirements at reduced temperatures, resulting in power comparable to the *spatial subsetting* scheme but provides a faster response time since our scheme does not power-off processor cores.

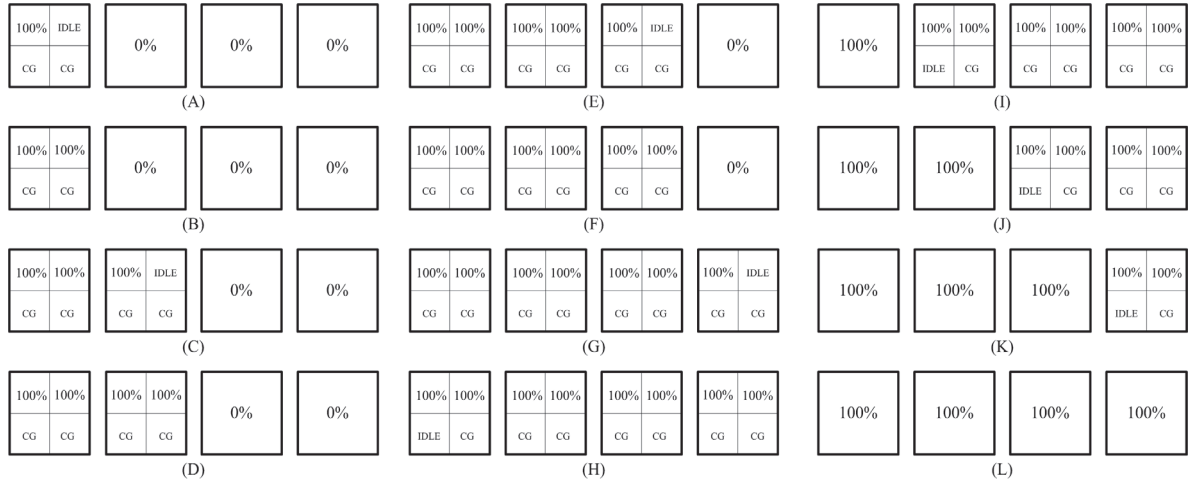


Figure 11: Proposed methodology across different utilization levels.

4 ASSIGNMENT OF WORLOAD BASED ON SLA

As an extension to the proposed methodology, we combine it with the G/G/m-model to reduce both the total power consumption and the response time degradation while meeting specific SLA requirements. Results from the queuing theory have been used to obtain measures like average execution velocity and average wait time to support capacity and workload planning of multi-processor systems for different workload variability (z). Using the approximation formulas for a G/G/m-model, we can reach an optimal agreement between high utilization of the processors (energy-conscious provisioning) and the target SLA requirements. For simplicity, consider a scenario where a specific workload requires 8 cores at 98% of utilization level, and assume that this workload can be linearly mapped to 9 and 10 cores at 87% and 78%, respectively as shown in Fig. 13. Fig. 14 shows the execution velocity for each of these 3 workload scenarios.

Execution velocity is the average ratio for the total amount of workload units that are served without any delay. The value ranges from 0 to 100 where the value 100 means that the workload does not encounter any wait delays for the system resources while the value 0 means that all work is delayed. Fig. 14 is derived using the formula given in (Müller-Colstermann, 2007). When setting the SLA for execution velocity of $>60\%$, using 10 processors to a utilization of 78% satisfies the requirement. On the other hand, using 8 processors result in an unacceptable execution velocity of 6.5%. Moreover, it is important to note that by increasing the number

of processors, there is no transition delay due to powering up a processor, and the only performance degradation results from the response delay of CG.

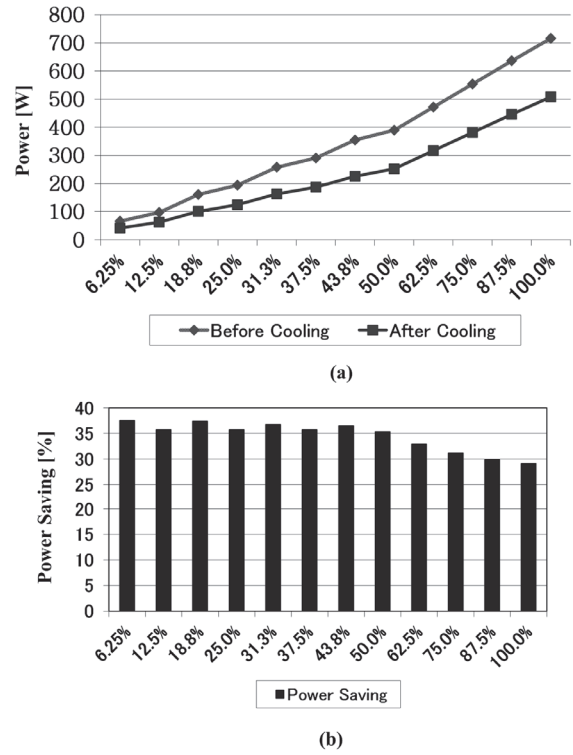


Figure 12: Power consumption at corresponding utilization levels of Fig. 10. Number in the figure represents the associated power saving from electronic cooling.

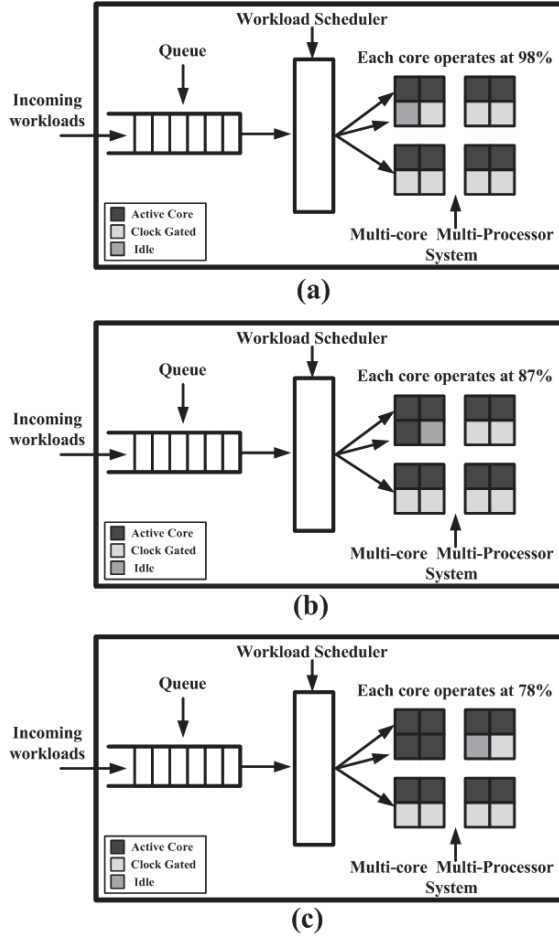


Figure 13: Required utilization level across different number of processors for the proposed methodology.

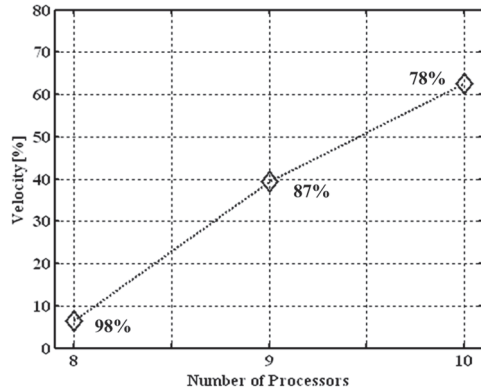


Figure 14: Execution velocity vs. number of processors. Number in the figure represents required utilization level.

Next, we evaluate the normalized average wait time, $E[W]$, for different values of workload

variability, z , where the normalization is performed with respect to the service time to the length of one unit. Here, workload variation represents the variation of request inter-arrival times and request sizes. We consider $0 \leq E[W] \leq 0.5$ for the good quality of service level. Similarly, notice how the system requires 10 processors at 78% of utilization to meet the average wait time requirements for $z \leq 5$ (see Fig.15).

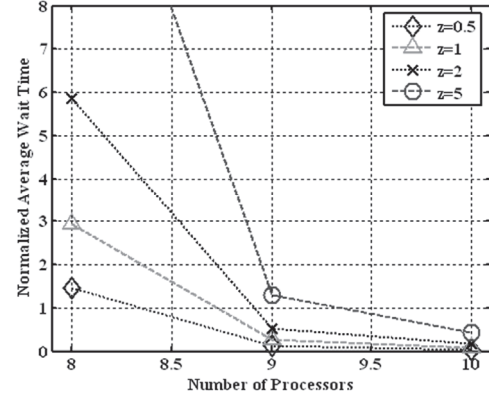


Figure 15: Normalized average wait time vs. number of processors as function of workload variability $z=0.5, 1.0, 2.0, 5.0$.

Finally, we summarize the results of our proposed methodology by comparing with *spatial subsetting*. The total amount of power consumption before and after cooling for the two schemes is shown in Fig. 16. As expected, the actively cooled system with the proposed methodology dissipates power that is comparable to the *spatial subsetting* scheme but enables superior response time for different levels of SLA. Analysis also shows that the overall system power savings of 35, 30, and 29% are obtained when using 8, 9, and 10 cores, respectively. It is worth noting that the amounts of saving decreases as we increase the number of cores as the cores now operate at lower utilization levels. Using a larger number of cores at lower utilization levels inevitably increases the total power consumption, but the system operates with much improved SLA. For instance, using 10 processors instead of 8 increase the total power consumption by 25%, but the system now operates at execution velocity of $>60\%$ and normalized wait time of ≤ 0.5 .

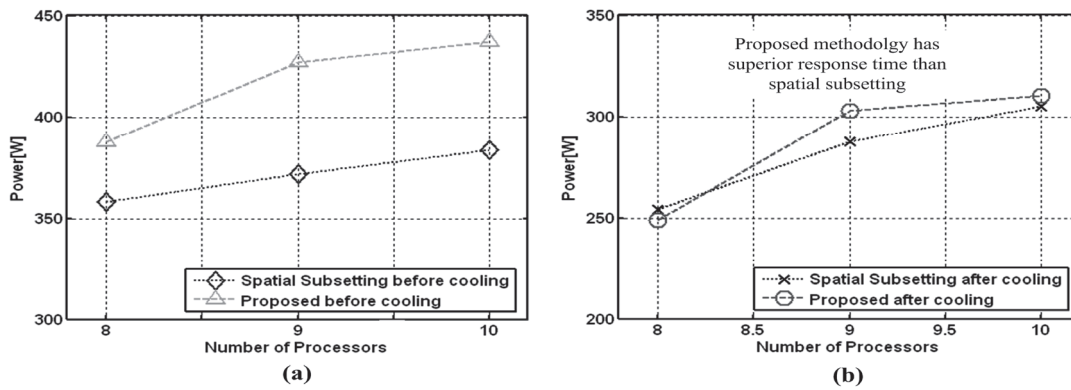


Figure 16: Power consumption vs. number of processors (a) before and (b) after cooling.

5 CONCLUSIONS

An energy-efficient workload scheduling methodology for HPC servers is presented using a highly efficient miniature scale refrigeration system for electronic cooling. By leveraging the benefits of clock gating at reduced temperatures, our proposed methodology results in total power consumption that is comparable to the *spatial subsetting* scheme. Moreover, it provides a response time of disabling clock gating which is orders of magnitude faster than waking from power gating or powering up a processor. Our actively cooled system results in $\geq 29\%$ power reduction over the non-refrigerated design across the entire range of utilization levels. Furthermore, combining our proposed methodology with the G/G/m-model, we show the trade-off between power and SLA requirements. Setting the target SLA requirement to execution velocity of $>60\%$ and normalized wait time of ≤ 0.5 , the number of required processors to execute a particular workload inevitably increased, leading to the 25% increase in total power consumption. Nevertheless, this still maintains 29% of power reduction, compared to non-cooled design.

While the results discussed in this paper can be directly applied to large-scale multi-server systems, overall system realization is still a big challenge and some important design issues of building such systems are overall power consumption, reliability, and cost. Furthermore, thorough understanding of the strong coupling between refrigerated server racks and CRAC units (cascaded cooling system) is needed for future research. Nevertheless, current data centers can consume up to 90% of the total energy from computer load and the energy required by the CRAC units. Decreasing the power dissipated or by the computer load is imperative as any

reduction in electronic heat can be compounded in the cooling system.

Finally, it would be interesting to explore different feedback-driven control solutions that provide capability to adapt to diverse environment, workload, and user constraints. This is relegated to future work. A model-based software framework that predicts and senses upcoming workloads and provides real-time information to refrigeration and electronic systems to tune compressor speed, temperature, and supply voltage are worthy of being studied in order to achieve optimal power performance.

REFERENCES

- U.S. EPA. Report to congress on server and data center energy efficiency. In *U.S. Environmental Protection Agency, Tech Report*, 2007.
- K. Rajamani, C. Lefurgy, J. Rubio, S. Ghiasi, H. Hanson, and T. Keller, "Power management for computer systems and data centers", Tutorial presented at the *2008 International and Symposium on Low Power Electronics and Design*, August, 2008.
- J. Humphreys and J. Scaramella, "The impact of power and cooling on data center infrastructure," Market Research Report, IDC, 2006.
- Tschudi, et al., "Data Centers and Energy Use – Let's Look at the Data", *ACEEE*, 2003.
- Lawrence Berkeley National Labs, Benchmarking: Data Centers, Dec 2007.
- M. Patterson, "The effect of data center temperature on energy efficiency," *Proceedings ITherm*, pp. 1167-1174, 2008.
- Jing, S., Ali, S., She, K., Zhong, Y., State-of-the-art Research Study for Green Cloud Computing. *Journal of Supercomputing*, Special Issue on Cloud Computing, 2011
- M. Lin, A. Wierman, L. Andrew, and E. Thereska,

- "Dynamic Right-Sizing for Power-Proportional Data Centers," *Proc. IEEE INFOCOM*, 2011.
- J. Luo, L. Rao, and X. L. Liu, "Data center energy cost minimization: a spatio-temporal scheduling approach," in *Proceedings of the INFOCOM 2013*.
- P. Chaparro, et al., "Understanding the Thermal Implications of Multicore Architectures," *IEEE Transactions on Parallel and Distributed Systems*, vol. 18, no. 8, pp. 1055-1065, August 2007.
- M. Ma, S. Gunther, B. Greiner, N. Wolff, C. Deutschle, and Tawfik Arabi, "Enhanced Thermal Management for Future Processors," *IEEE Symposium on VLSI Circuits of Technical Papers*, pp. 201-204, June 2003.
- J. Tschanz, S. Narendra, Y. Ye, B. Bloechel, S. Borkar, and V. De, "Dynamic Sleep Transistor and Body Bias for Active Leakage Power Control of Microprocessors," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 11, pp. 1838-1845, November 2003.
- D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: A framework for architectural-level power analysis and optimizations," in *International Symposium on Computer Architecture*, June 2000.
- T. Sato and T. Funaki, "Power-Performance Trade-off of a Dependable Multicore Processor," in 13th Pacific Rim International Symposium on Dependable Computing (PRDC), 2007.
- R. Ghosh, V. K. Naik, and K. S. Trivedi, "Power-Performance Trade-offs in IaaS Cloud: A Scalable Analytic Approach," in IEEE/IFIP DSN Workshop on Dependability of Clouds, Data Centers and Virtual Computing Environments (DCDV), 2011.
- E. Pinheiro, R. Bianchini, E. V. Carrera, and T. Heath, "Load balancing and unbalancing for power and performance in cluster-based systems," *Workshop on Compilers and Operating Systems for Low Power*, 2001.
- J. Chase, D. Anderson, P. Thakur, and A. Vahdat, "Managing Energy and Server Resources in Hosting Centers," *Proceedings of the 18th Symposium on Operating systems Principles SOSP'01*, October 2001.
- Y. Chen, A. Das, W. Qin, A. Sivasubramaniam, J. Srebric, Q. Wang, and J. Lee, "Managing Server Energy and Operational Costs in Hosting Centers," *SIGMETRICS Performance Evaluation Review*, vol. 33, no. 1, pp. 303-314, 2005.
- F. Ahmad and T. Vijaykumar, "Joint optimization of idle and cooling power in data centers while maintaining response time," *Architectural Support for Programming Languages and Operating Systems*, 2010.
- B. Müller-Clostermann, "Using G/G/m-Models for Multi-Server and Mainframe Capacity Planning," *ICB Research Report*, no. 16, May 2007.
- R. McGowen, C. A. Poirier, C. Bostak, J. Igonowski, M. Millican, W. H. Parks, and S. Naffziger, "Power and temperature control on a 90-nm Itanium family processor," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 1, pp. 228-236, January 2006.
- T. D. Burd, T. A. Pering, A. J. Stratakos, and R. W. Brodersen, "A Dynamic Voltage Scaled Microprocessor System," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 11, pp. 1571-1580, November 2000.
- K. J. Nowka, G. D. Carpenter, E. W. MacDonald, H. C. Ngo, B. C. Brock, K. I. Ishii, T. Y. Nguyen, and J. L. Burns, "A 32-bit PowerPC System-on-a-Chip With Support for Dynamic Voltage Scaling and Dynamic Frequency Scaling," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 11, November 2002.
- S. Heo, K. Barr, and K. Asanovic, "Reducing power density through activity migration," *International Symposium on Low Power Electronics and Design*, Aug. 2003.
- K. Zhang et al., "SRAM design on 65-nm CMOS technology with dynamic sleep transistor for leakage reduction," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 4, April 2005, pp. 895-901.
- S. Henzler, T. Nirschl, S. Skiathitis, J. Berthold, J. Fischer, P. Teichmann, F. Bauer, G. Georgakos, and D. Schmitt-Landsiedel, "Sleep transistor circuits for fine-grained power switch-off with short power-down times," in *Proc. Int. Solid-State Circuits Conf.*, 2005, pp. 302-303.
- D. Copeland, "64-bit Server Cooling Requirements," *IEEE SEMI-THERM Symposium*, 2005.
- R. Mahajan, C. P. Chiu, and G. Ghysler, "Cooling a Microprocessor Chip," in *Proceedings of the IEEE*, vol. 94, no. 8, Aug. 2006.
- A. G. Agwu Nnanna, "Application of refrigeration system in electronics cooling," *Applied Thermal Engineering*, vol. 26, pp. 18-27, 2006.
- R. C. Chu, R. E. Simons, M. J. Ellsworth, R. R. Schmidt, and V. Cozzolino, "Review of Cooling Technologies for Computer Products," *IEEE Trans. on Device and Material Reliability*, vol. 4, no. 4, pp. 568-585, Dec. 2004.
- P. E. Phelan, V. A. Chiriac, and T. T. Lee, "Current and Future Miniature Cooling Technologies for High Power Microelectronics," *IEEE Trans. on Components and Packaging Technologies*, vol. 25, no. 3, pp. 356-365, Sep. 2002.
- S. Trutassanawin, E. Groll, V. Garimella, and L. Cremaschi, "Experimental Investigation of a Miniature-Scale Refrigeration System for Electronics Cooling," *IEEE Trans. on Components and Packaging Technologies*, vol. 29, no. 3, pp. 678-687, Sep. 2006.
- D. M. Carson, D. C. Sullivan, R. E. Bach, and D. R. Resnick, "The ETA-10 liquid-nitrogen-cooled supercomputer system," *IEEE Trans. Electron. Devices*, vol. 36, no. 8, pp. 1404-1413, Aug. 1989.
- Won Ho Park, Tamer Ali, and C.K. Ken Yang, "Analysis of Refrigeration Requirements of Digital Processors in Sub-ambient Temperatures," *Journal of Microelectronics and Electronic Packaging*, vol. 7, no. 4, 4th Qtr 2010.
- Won Ho Park and C.K. Ken Yang, "Effects of Using Advanced Cooling Systems on the Overall Power Consumption of Processors," accepted for *IEEE*

Transactions on Very Large Scale Integration Systems.

- I. Aller et al., "CMOS Circuit Technology for Sub-Ambient Temperature Operation," Proc. Int. Solid-State Circuits Conf., 2000, pp. 214-215, Feb. 2000.
- J. Rabaey, A. Chandrakasan, and B. Nikolic, Digital Integrated Circuits: A Design Perspective; 2nd ed., 2003.
- A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-Power CMOS Digital Design," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 4, pp. 473-484, April 1992.
- R. Kumar, K. Farkas, N.P. Jouppi, P. Ranganathan, and D.M. Tullsen, "Single-ISA Heterogeneous Multi-Core Architectures: The Potential for Processor Power Reduction," *Proc. Int'l Symp. Microarchitecture*, Dec. 2003.
- N. A. Kurd, J. S. Barkatullah, R. O. Dizon, T. D. Fletcher, and P. D. Madland, "A multigigahertz clocking scheme for Pentium 4 microprocessor," *IEEE Journal of Solid-State Circuits*, vol. 36, pp. 1647-1653, November 2001.