UNIVERSITY OF CALIFORNIA

Los Angeles

Effects of Cooling on the Overall Power Consumption of Processors

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical Engineering

by

Won Ho Park

2012

ABSTRACT OF THE DISSERTATION

Effects of Cooling on the Overall Power Consumption of Processors

by

Won Ho Park

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2012

Professor Chih-Kong Ken Yang, Chair

To address the power and thermal problems of high-performance computing systems, the possibility of using advance cooling systems is explored to realize overall system power improvement that includes the cost of cooling power. Realistic system-level modeling that includes the electronic and refrigeration systems will expedite the analysis of the optimal operating temperature points, the amount of total power reduction at reduced temperatures, and the dependence of the refrigerated performance on the power profile of the electronics. Highly-efficient miniature-scale refrigeration system for electronic cooling is developed to experimentally demonstrate the amount of total power saving for processors at various operating conditions. A processor that dissipates 175.4W of maximum power with 30% electronic leakage power operating at 97°C is cooled using our refrigeration system. Measurements show that with a minimum refrigeration COP of 2.7, the processor operates with junction temperature <40°C and offers a 25% total system power reduction over the non-refrigerated design. Not only the measurement results validate our system-level model, but this experiment is the *first demonstration* of active cooling that lead to reduced total wall power. Furthermore, a model that captures different relations and parameters of multi-core processor and the refrigeration system is constructed based on the measured results. This model is used to present an energy-efficient workload scheduling that optimizes power efficiency under the actively cooled environment that can potentially be applied to large scale multi-core, multi-processor computing environment. Finally the proposed methodology is combined with the G/G/m-models to reduce both total power and response time degradation while meeting target SLA requirements.

The dissertation of Won Ho Park is approved.

Youngho Ju

Jason Woo

Dejan Markovic

Chih-Kong Ken Yang, Committee Chair

University of California, Los Angeles

2012

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisor, Professor Chih-Kong Ken Yang, for his continuous guidance and patience throughout the course of my research and thesis.

VITA

2005	B.S. in Electrical Engineering Korea University Seoul, Korea
2008	M.S. in Electrical Engineering University of California, Los Angeles Los Angeles, California, USA
2006-2012	Graduate Student Researcher High-Performance Mixed-Mode Circuit Design Group Department of Electrical Engineering University of California, Los Angeles Los Angeles, California, USA
Summer 2006	Intern, Samsung Advanced Institute of Technology, Inc, Kihung, Korea
Summer 2008	Intern, Samsung Semiconductors, Inc, Hwasung, Korea
Summer 2009 & 2010	Intern, Texas Instruments, Dallas, Texas, USA
2011 - Present	Intern, Pluribus Networks, Palo Alto, California, USA

PUBLICATIONS

Won Ho Park, Tamer Ali, 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, C.K. Ken Yang, “Analysis of Total System Power Efficiency Using Embedded Thermoelectric Coolers,” International Microsystems, Packaging, Assembly, and Circuits Technology Conference, 2011

Won Ho Park, C.K. Ken Yang, “Total System Power Reduction Using a Refrigeration System for Electronic Cooling,” International Microsystems, Packaging, Assembly, and Circuits Technology Conference, 2011

Won Ho Park, C.K. Ken Yang, “Effects of Using Advanced Cooling Systems on the Overall Power Consumption of Processors,” in revision for IEEE Transactions on Very Large Scale Integration

Won Ho Park, C.K. Ken Yang, “Effects of Cooling on Workload Management in High Performance Computing Processors,” in revision for IEEE Transactions on Computers

CHAPTER 1

Introduction

Aggressive CMOS technology scaling over the last few decades led to the tremendous growth in the semiconductor industry. The ability to integrate billions of transistors of into a single chip increased computational capacity. However, besides the traditional performance constraint, power dissipation and thermal issues have been identified as key limitations for the design of high performance computing (HPC) microelectronic chips. The limitations are due to increasing power density with technology scaling, and operating processors for HPC applications near maximum power limit in order to maximize server performance. As a result, reducing the operating temperature using refrigeration systems for electronic cooling has attracted recent interest as a possible option that offers solutions to both the power and thermal problems.

This chapter begins with describing some trends in system power dissipation as motivation behind this dissertation. We show the problem of power consumption in current data centers and servers and how reducing operating temperature of HPC microprocessors can potentially address the problem. This dissertation focuses on using a refrigeration system for electronic cooling can not only be used to improve performance but also to realize overall power improvement that includes the cost of cooling power. The outline of the thesis is presented in the second section.

1.1 Motivation

Increase in energy consumption due to the tremendous growth in the number, size, and uses 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 U.S. 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 is a 60% increase since year 2006 [1]. Worldwide trend of energy consumption in data centers tracks with the trend of the U.S [2].

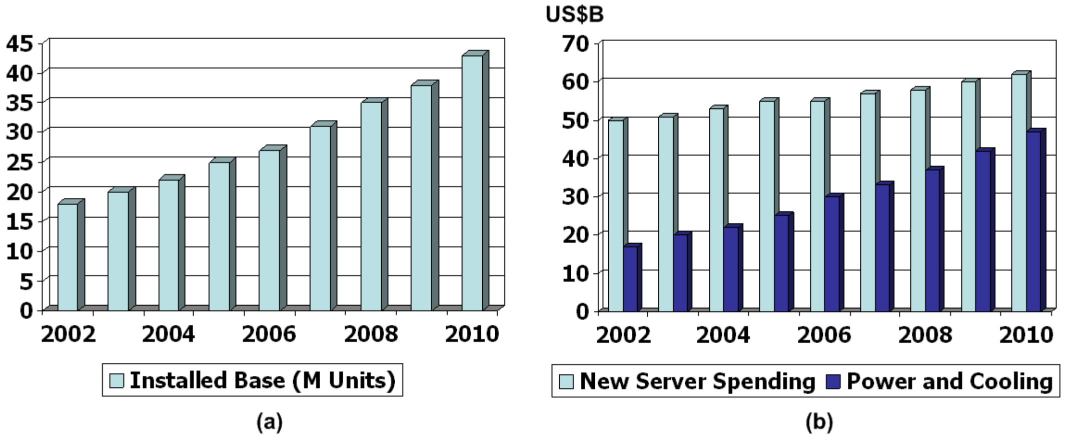


Figure 1-1: (a) Number of worldwide installed bases of data center. (b) Worldwide spending on new servers and operation cost.

Figure 1.1 shows the number of installed bases of data centers, worldwide new server spending, and electric power and cooling costs. Despite the dramatic 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 more dense, power density has been increasing by approximately 15% annually [3], raising the electric energy consumption for operating its servers and cooling system. This energy consumption is reflected in the electricity bill, where the Figure 1.1.b shows the amount of electric bill consumed by data centers. In fact, the cost has been increasing by approximately $4 billion per year. If this problem is not properly addressed, it is evident that IT operation cost will soon outweigh the installation cost.

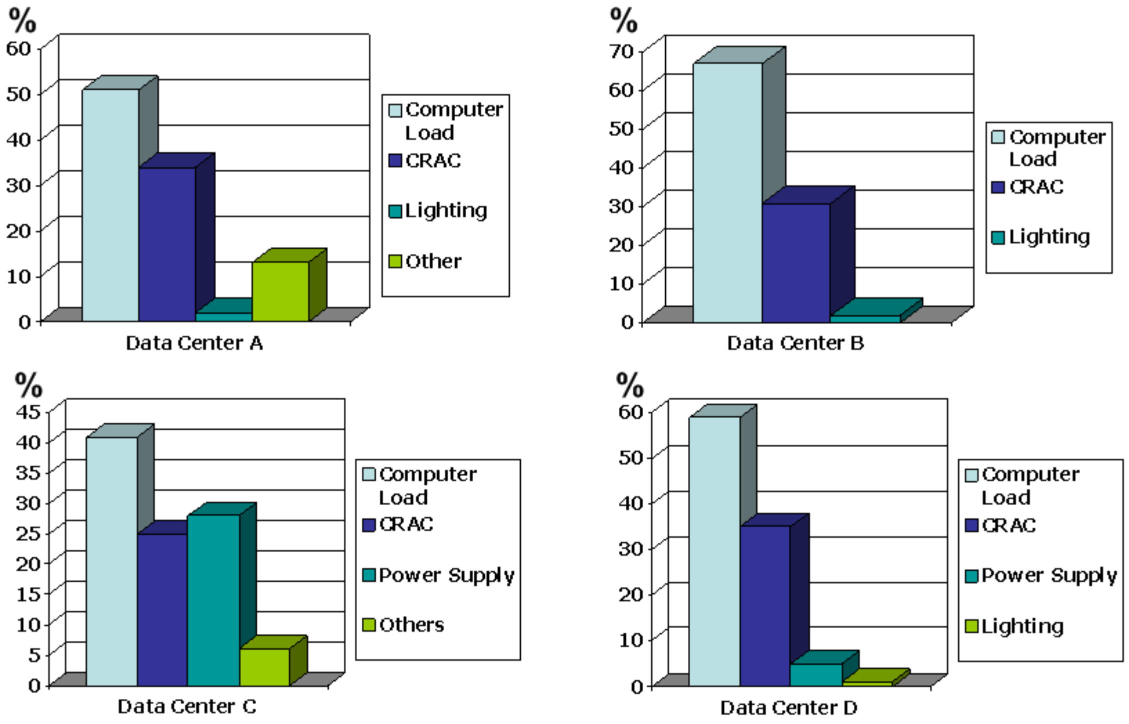


Figure 1-2: Energy breakdown of different data centers.

Detailed energy breakdown of different types of data centers is shown in Figure 1.2 [2], [4] - [6]. 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 **C**omputer **R**oom **A**ir **C**onditioner (CRAC) unit, which removes the heat generated by the computer load. Therefore, reducing energy of both the computing and cooling is imperative in minimizing energy usage of data centers. Furthermore, there exists a strong coupling between the energy consumed by the computer load and 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 Celsius [2].

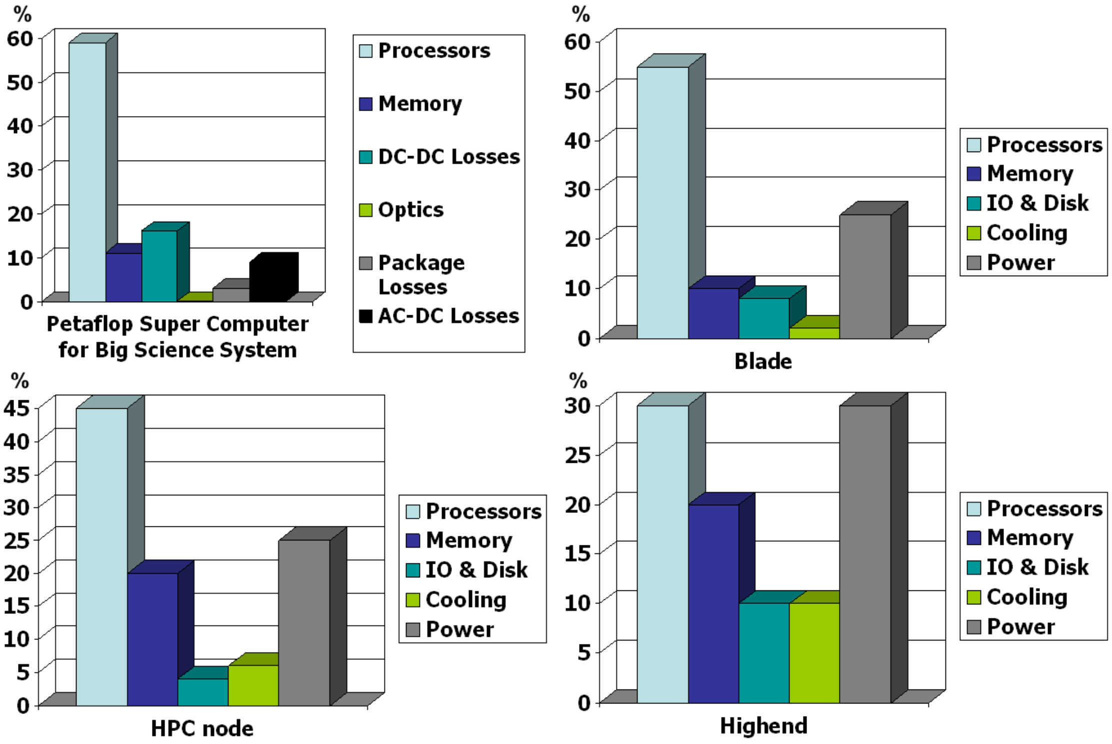


Figure 1-3: Power breakdown of different types of servers.

Data centers consists of hundreds or even thousands of server racks where each rack can draw more than 20kW of power; inability to meet such enormous power requirements have resulted in over 40% of customers reporting that the power demand is outstripping their power supply [3]. This situation creates an unprecedented demands for minimizing power consumption of high performance servers. Consequently, optimizing energy usage of multi-server systems is an interest across multiple disciplines, with the energy efficiency being a major research interest.

The focus of most efforts on energy/power saving in multi-server systems is on processor elements [7] - [10], [44]. Power breakdown on different components of various server types is shown in Figure 1.3 [2]. The distribution of power consumption varies with the server class and also depends on the configuration and usage. However, it is clear that processors are the dominant source of power consumption, reaching up to 60% for petaflop supercomputer node designs.

Approaches for energy saving often adopt both the software-based energy-aware workload scheduling 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. Often the software utilizes special hardware supports offered by HPC processors, such as stopping a processor, dynamic voltage frequency scaling [11] - [13], thread migration [14]. Although these special hardware support can be used to reduce power consumption, these techniques often require some transition time in and out of different power states and imply some performance degradation.

Along with these techniques, actively cooling the processors using refrigeration systems have attracted recent interest as a practical option to ease the power problems in HPC processors for sever type applications [15] - [17]. The main benefit of lowering temperature is to allow the system to maintain a low junction temperature while operating with high performance. Often, active cooling results in increased overall power dissipation due to the combined electronic power and refrigeration power.

Although reduced temperature operation of CMOS circuits improve the electronic performance, there has been little work that synthesizes the cooling power consumption in order to quantify the overall benefit to the system. For instance, the improved speed performance can be trade-off into a power reduction with constant performance by reducing the supply voltage, V_{dd}. The research in this dissertation is toward the analysis and design of electronic system with refrigerated system for electronic cooling to investigate whether the system offers an overall power reduction over the non-cooled design. Particularly, we focus on two types of electronic cooling systems: embedded thermoelectric coolers and vapor compression refrigeration system. In order to achieve this goal, a realistic system-level model that includes both the electronic and the refrigeration systems is introduced. The model is used to explore the optimal operating temperatures and the amount of total power reduction at reduced temperatures.

Developing a highly-efficient miniature-scale refrigeration system for electronic cooling is one key element of this research. We have developed a refrigeration system that operates at low temperatures with high efficiency. This refrigeration system is used to experimentally demonstrate the amount of total power saving for processors at various operating conditions. Based on the measured results, a model that captures different relations and parameters of multi-core processor and the refrigeration system is presented. Finally, this model is extended to illustrate the results of optimizing the power of multi-core multi-processor systems under the actively cooled environment and to investigate different methodologies of workload scheduling to maximize power efficiency while minimizing performance degradation. Results presented in this dissertation suggest that there exists an efficient methodology under the actively cooled environment that optimizes power efficiency.

1.2 Organization

The dissertation starts with an introductory chapter and is then divided into four parts that deal with system modeling, analysis, implementation, and different power minimizing techniques. Chapter 2 describes the mechanisms for power dissipation of digital CMOS ICs and the benefits of reduced temperature operation. The chapter also introduces a performance metric of two different types of refrigeration systems. The models of such cooling systems are combined with the electronics model and are discussed in Section 2.1 and 2.2, respectively. The basic power saving techniques of multi-core multi-processor systems are also briefly discussed in Section 2.3.

Chapter 3 includes the analysis to support the feasibility of using vapor compression refrigeration system for electronic cooling. We explore the impact on overall system power, desired operating temperatures, and the dependence of the refrigerated performance on the power profile of the electronics. In Section 3.1, the benefits of microprocessors at sub-ambient temperatures are discussed in a case study. Dependence on the characteristics of the electronics and refrigeration systems is presented in Section 3.2 and 3.3, respectively. From these results, performance requirements for the refrigeration are discussed for different electronic power-dissipation characteristics.

Chapter 4 explores the amount of total power reduction using embedded thermoelectric coolers (eTECs). Recent advances in eTECs provide possibility of mitigating thermal problems in HPC systems. The effects of localized cooling on different functional blocks are explored using a model that incorporates both a real-world microprocessor and thermoelectric cooling systems as discussed in Section 2.2. Our analysis in Section 4.1 indicates that an optimal operating point exists and depends on the parameters of electronics and cooling systems. In Section 4.2, we provide our simulation results of localized spot cooling various blocks of a high performance microprocessor.

Our discussion on system level modeling and analysis ends in Chapter 4. While our analysis in Chapter 3 provides the analysis to support the feasibility of using refrigeration systems for electronic cooling in terms of overall power performance, obtaining this goal requires a highly-efficient refrigeration system. In Chapter 5, a refrigeration system is developed and experimentally tested to demonstrate that cooling the high performance microprocessor can indeed lead to overall system power improvement. Overall system description and our experimental setup is discussed in Section 5.1. We go on to describe the performance of the refrigeration system in Section 5.2. Measurement results validate our analysis of Chapter 3 and is discussed in Section 5.3. The chapter then concludes with a discussion of future refrigeration system for electronic cooling.

Chapter 6 explores an energy-efficient workload scheduling methodology for multi-core multi-processor systems under the actively cooled environment that improves the overall system power performance with minimal response time degradation. Using our miniature-scale refrigeration system from Chapter 5, 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. In Section 6.1, the measured results are used to establish a power model for a multi-core processor based on varying levels of activity under an actively cooled environment. Section 6.2 applies various workload scheduling algorithms to this model to explore the best approach to maximize energy efficiency. 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, and is discussed in Section 6.3. Finally, Chapter 7 concludes this dissertation with discussions for future work.

CHAPTER 2

Background

Several key power and technology attributes for semiconductor ICs from the 2009 International Technology Roadmap for Semiconductors (ITRS) are listed in Table 2.1 [18]. The power dissipation for single-chip packages in high-performance systems is predicted to rise to 161W in 2011 while the maximum junction temperature is restricted to 85ºC or lower. If temperature and power are not properly managed, both the chip and package level performance and reliability are critically affected. In order to meet both requirements, the overall junction-to-ambient thermal resistance needs to decrease. Table 2.1 shows that the required overall junction to ambient thermal resistance for high-performance ICs for ambient temperature of 45°C must be substantially less than 0.3°C/W in the near future.

The use of cooling through active heat removal is being evaluated as a viable option that might offer solutions to both the power and thermal problems [19] - [21]. Such cooling leads to lower junction temperatures below the common forced-air cooling which in turn can alleviate thermally induced reliability issues. Cooling can also reduce power consumption and/or boost the speed performance of high-end microprocessors. Various refrigeration technologies for electronic cooling such as thermoelectric, cryogenics, and vapor compression with the associated advantages and disadvantages have been reported in [19] - [21]. Recently, two types of active cooling have been commonly discussed: eTECs [23] - [25] and highly-efficient vapor compression refrigeration systems [21] - [22]. The two different cooling technologies can be distinguished in how they can be applied in a system. Embedded TECs have the potential to locally cool different portions within an IC (*localized spot cooling*). Refrigerated cooling in which the evaporator directly attaches to the processor package is often considered in the context of cooling the entire dice (*chip-level cooling*). This type of refrigerated systems can possibly be extended to directly cool multiple dice in multiple chassis depending on the size of the compressor. Thereby, the potential of power optimization of multi-core multi-processor systems under the cooled environment using miniature-scale refrigeration system can be investigated based on different workload scheduling methodologies.

This chapter focuses on combining an electronics model with a realistic active-cooling model of two different types of refrigeration systems: eTECs and vapor compression refrigeration systems. For a given set of parameters, the model can be used to determine the system power at various operating temperatures and power. Section 2.1 discusses the system model and reviews the well-known parameters for the electronic and vapor compression refrigeration systems. The model is used in the analysis in Chapter 3. A realistic system-level model that includes both the electronic characteristics of various functional blocks and the realistic performance of eTECs is introduced in Section 2.2. This model is used in the analysis of Chapter 4. Lastly, the model from Section 2.1 is extended to capture different relationships and parameters of our multi-core processor and the refrigeration system. Such model can be used to illustrate the potential of power optimization of multi-core multi-processor systems and investigate different methodologies of workload scheduling which is discussed in Chapter 6. Basic power saving techniques and power-aware workload algorithms in multi-processor server systems are discussed in Section 2.3.

Year	2008	2009	2010	2011
Gate Length (nm)	29	27	24	22
Allowable Maximum Power (W)	146	143	146	161
Power Density (W/mm^2)	0.47	0.46	0.47	0.52
Maximum Junction Temperature (°C)	85	85	85	85
Required Thermal Resistance (°C/W)	0.27	0.28	0.27	0.25
$I_{d,sat}$: NMOS Drive Current (uA/um)	1006	1317	1370	1333
$I_{d,leak}$: NMOS Drive Current (uA/um)	0.13	0.17	0.46	0.71
Full Chip Leakage[1]	1.5	2	2.5	2.75

Table 2.1 International Technology Roadmap for Semiconductors 2008. ([1]:Normalized to Full-Chip leakage power dissipation in 2007)

2.1 Electronic System with Vapor Compression Refrigeration

Figure 2.1 illustrates three possible configurations of a computing server unit. Figure 2.1.a is a common air-cooled system with a fan forcing air flow over all the electronics. The best-case scenario for air cooling the microprocessor is shown in Figure. 2.1.b where the microprocessor heat-sink is directly cooled with air from the ambient temperature. The corresponding refrigerated-cooling system would directly cool the heat-sink as shown in Figure. 2.1.c. A fan is included in the system to cool the condenser.

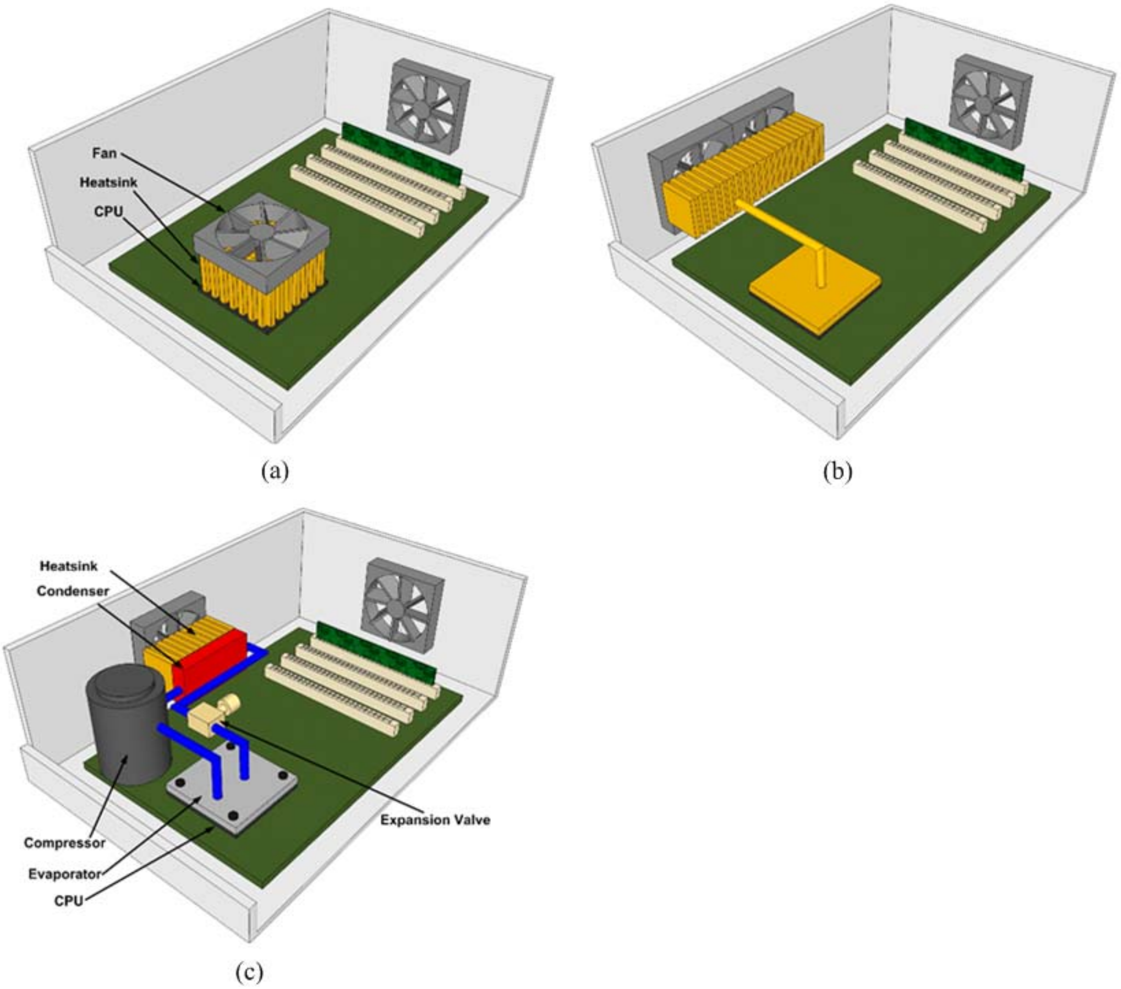


Figure 2-1: Three possible configurations of a computing server unit with (a) a typical configuration of a fan for the chassis, (b) dedicated fan-cooling for the processor, and (c) a refrigerated-cooling for the processor.

The power and heat model is illustrated in Figure. 2.2. A digital processor is the heat load consuming power, $P_{electric}$. In Figure 2.2.a, the heat load is maintained at $T_{junction0}$ by a forced-convection air cooler, which rejects $P_{electric}$ to $T_{ambient}$ through the junction-to-case thermal resistance (R_{jc}) and case-to-ambient thermal resistance (R_{ca}). The associated fan power needed to obtain R_{ca} is denoted by P_{fan1}. With a given thermal resistance, the junction temperature of the system is determined by the total power consumption and the ambient temperature. The total power dissipation of this configuration, which is denoted by P_{total_air}, is given by the sum of $P_{electric}$ and P_{fan1}. Figure 2.2.b shows the model of the same system with a refrigeration system. In this model, the heat load is being maintained at $T_{junction}$ ($<T_{junction0}$). The cooling process is performed between hot-end temperature of the condenser (T_{hot}) and cold-end temperature of the evaporator (T_{cold}) to allow the transfer of $P_{electric}$ to the ambient environment. The required power for refrigeration is represented by $P_{refrigeration}$. Also the thermal resistance of refrigeration heatsink and its required fan power are modeled by R_{ra} and P_{fan2}, respectively. Here, the total power dissipation, P_{total_ref}, is obtained by the sum of $P_{electric}$, $P_{refrigeration}$, and P_{fan2}. Various sources of power dissipation in the two systems are summarized. The following sub-sections review the relationship between power dissipation and temperature for electronics and for refrigeration. These relationships and the parameters that characterize the performance of the electronics and refrigeration are used in a model built in MATLAB. The model is used for the analysis and design in Chapters 3 and 5.

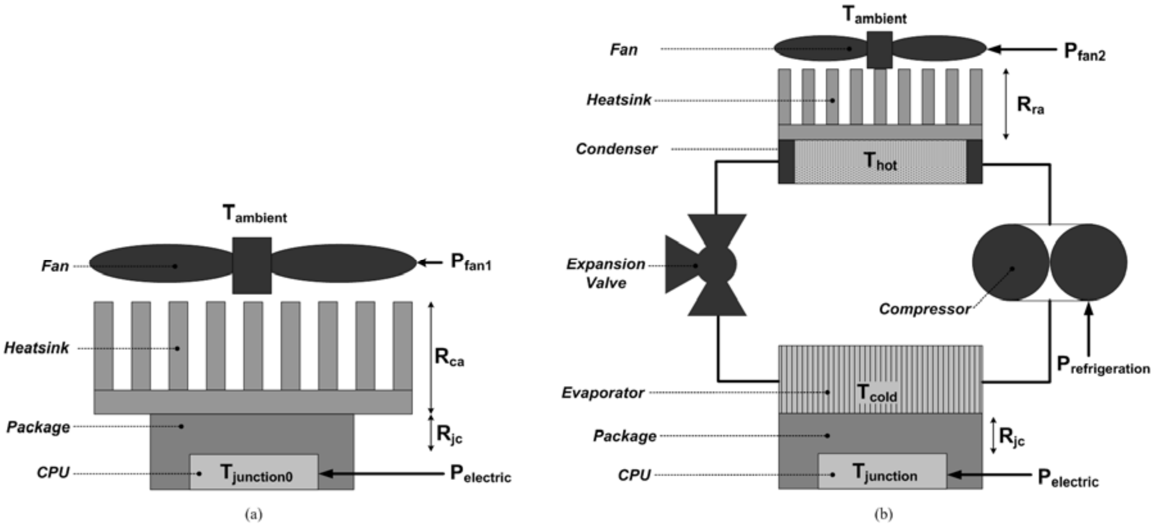


Figure 2-2: (a) Model of conventional air cooler. (b) Model of the same electric system using a refrigeration system.

2.1.1 Sources of Electronic Power

The total electronic power dissipation of a microprocessor can be approximated by the sum of active power and subthreshold leakage power [26], [58].

$$P_{electronic} \approx P_{active} + P_{leakage} \qquad (2.1)$$

$$P_{active} = \alpha * C_{switched} * f_{clk} * V_{dd}^{\ 2} \qquad (2.2)$$

$$P_{leakage} = V_{dd} * I_{leakage} = V_{dd} * I_0 \exp\left(\frac{-V_{th}}{kT_{junction}/q} \right) \qquad (2.3)$$

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 nanometer-scale technologies, the switches that implement the logic result in a leakage current to flow for 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 relationship to the ratio of a transistor's ON/OFF threshold, V_{th}, and the thermal voltage, $kT_{junction}/q$. The $P_{leakage}$ equation simplifies the dependences 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 [58]. The delay of logic switching is proportional to the switch and interconnection resistances and capacitances. Equation 2.4 shows this relationship for the delay of a logic gate. The current is a function of temperature and primarily depends on the carrier mobility (μ ~T^{-1}). A designer can typically trade-off any improved speed performance into a power reduction with constant performance by reducing the supply voltage, V_{dd}.

$$\text{Delay} \propto RC \propto \frac{V_{dd}}{I_{logic}} \qquad (2.4)$$

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 lower V_{th}. Figure 2.3 graphs the change in performance with temperature of a 65nm technology device. Figure 2.3 also shows the reduction in active and leakage electronic power with constant performance.

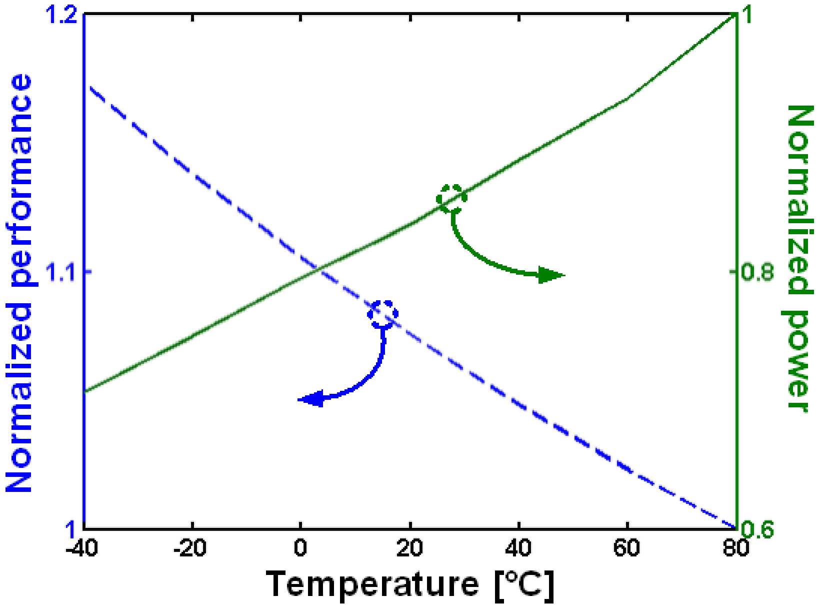


Figure 2-3: Normalized performance and power as a function of temperature of a 65nm technology device.

The percentage of electronic power due to its leakage and active components depends on a number of factors. The transistor technology can determine the $P_{leakage}/P_{electric}$ ratio at the level of the logic gates. Table 2.1 shows the ratio for an NMOS transistor for several different nanometer-scale technologies. With technology scaling, the supply voltage lowers to maintain device reliability and constrain power consumption. To compensate for this performance loss, V_{th} scales down correspondingly. Such scaling results in a dramatic increase of leakage power since leakage power depends exponentially on V_{th}. The leakage power is further exacerbated with technology scaling because of such problems as drain-induced barrier lowering (DIBL) and short channel effect.

The functionality of the digital block impacts $P_{leakage}/P_{electric}$ ratio through the average activity factor, α. For instance, memory blocks have considerably different α compared to arithmetic units. Examples of the ratio are shown Table 2.1.a for several functional blocks. The implementation of a functional block also has impact through the activity factor. As an example, a highly parallelized design has a tendency to have a higher ratio than the one that is not parallelized. Different applications and architectures such as microprocessors and digital signal processors also have different ratios as shown in the bottom entries in Table 2.1.b. The ITRS roadmap projects (Table 2.1) that full-chip leakage power dissipation increases with device scaling. Specifically, analysis of microprocessor power consumption [27] shows average leakage power ratio increases with device scaling, reaching up to 50% from the 65nm technology generation and beyond. Finally, the ratio in the table is typically an average value. The ratio changes dynamically depending on the instantaneous activity of the digital block. Leakage can dominate when the system is in stand-by, whereas active power is typically dominant with high-levels of logic activity.

As seen above, leakage power becomes important for high performance ICs. Due to its exponential dependence on temperature, $T_{junction}$ in turn, is determined by the total power consumption and system packaging technology. As a result of this strong coupling between $T_{junction}$ and leakage power, temperature-aware power modeling becomes necessary for appropriate thermal management and accurate performance optimization/estimation.

Functional Unit	Leakage Percentage (%)
IntExec	24
Bpred	44
L2 Cache	53

(a)

Application		Leakage Percentage (%)
DSP [29]		48
DSP [30]		25
SRAM [31]		53
Processor [32]	@ Average Power	36
16 Core Processor [33]	@ Maximum Power	12
Processor Leakage Trend [27]		50

(b)

Table 2.2: Leakage percentage (a) for various functional blocks within a microprocessor [28] and (b) for various types of integrated circuits.

2.1.2 Cooling Power and Refrigeration Efficiency

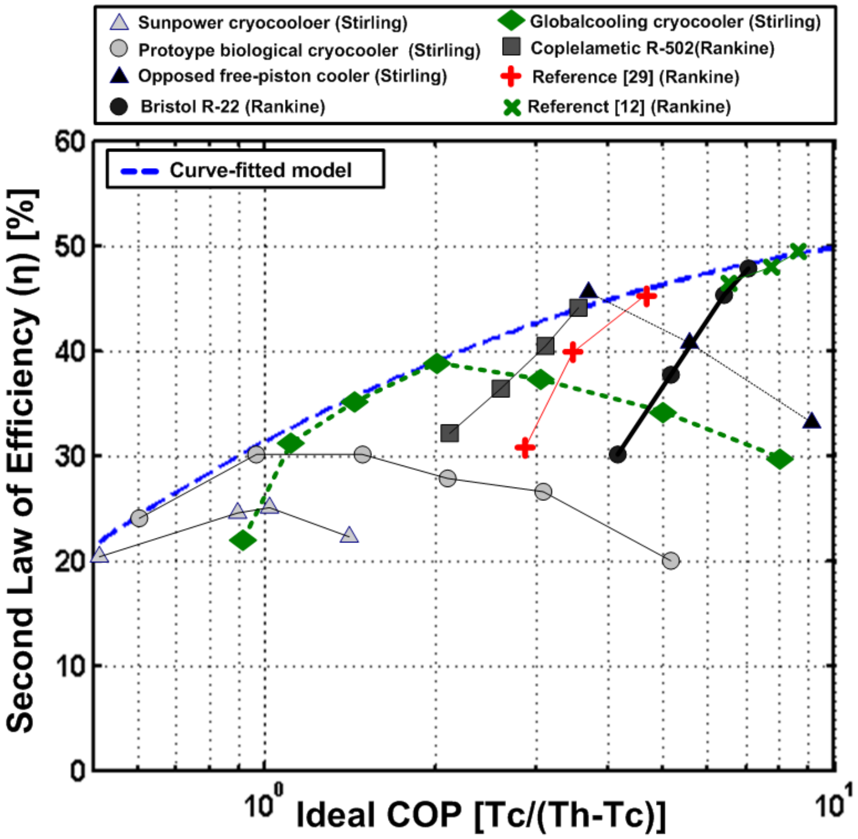


Figure 2-4: The performance of different refrigeration systems.

The power cost of cooling to a desired junction temperature includes the power cost of the refrigeration and fans and the thermal resistances [34] - [35]. Refrigeration power which accounts for the Carnot cycle and compressor power is modeled using the coefficient-of-performance (COP_{real}). This work uses the COP data from [21], [36] - [39]. The refrigeration power can be expressed in terms of the COP_{real} and the COP of the Carnot cycle with the well-known Equations (2.5) and (2.6) where T_{cold} is the cold-end temperature of the evaporator (<$T_{junction}$), T_{hot} is the temperature at the condenser (>$T_{ambient}$), η is the second-law efficiency, and $P_{electric}$ is the power dissipation of the heat source [57].

$$\frac{P_{electronic}}{P_{refrigeration}} = COP_{real} = \eta COP_{carnot} = \eta \frac{T_{cold}}{T_{hot} - T_{cold}} \quad (2.5)$$

$$P_{refrigeration} = \frac{P_{elctronic}}{\eta} \cdot \left(\frac{T_{hot}}{T_{cold}} - 1 \right) \quad (2.6)$$

Using the COP data, Figure 2.4 shows the efficiency data of different refrigeration products at different temperature ranges. Each system is optimally designed to function in the temperature ranges of various commercial applications. The COP of each system degrades as the temperature deviates from the optimal operating point. Since the analyses in Chapter 3 explore the optimal temperature across a broad temperature range, instead of using a single refrigeration system, a curve is fitted based on the optimal efficiency points across a number of the refrigeration products spanning the temperature range. In other words, this curve is a realistic upper bound for the achievable refrigeration efficiency at each temperature.

While it is difficult to simulate a complete system that perfectly matches the real environment and circuit/system impairments, the methodology presented above allows a quick yet accurate simulation that reflects the impact on different parameters on the overall system performance. Our model serves as a useful tool to evaluate overall system performance at different operating conditions. Performance requirements for the refrigeration systems can also be discussed for different electronic characteristics. The results enable a designer to best choose or design the cooling system for the application. The overall system power consumption that includes the cost of cooling power is used to illustrate the performance comparison throughout this dissertation.

2.2 Electronic System with Embedded Thermoelectric Coolers

Since cooling using vapor compression refrigeration system is performed at the chip level, a single temperature setting for the entire IC may not result in the optimum operating point for individual functional blocks within the integrated circuit. This is because each functional block within a processor can have considerably different power dissipation and power density. Integrated eTECs for localized cooling have been discussed and shown in [23] - [25]. A thermoelectric cooler (Figure 2.5) is a small electronic heat pump that has the advantage of no moving parts and silent operation. Thermoelectric cooling uses the *Peltier* effect to create a heat flux between the junctions of dissimilar materials. By adjusting DC current, heat flow can be proportionally modified allowing precise control of the junction temperature. The movement of heat provides a cooling capability that is well-suited for applications where temperature stabilization or direct cooling is required. Cooling research has been investigating using eTEC to locally and selectively cool an area with high power density to resolve thermal problems in integrated circuits. It is well known that cooling heat flux of a TEC is inversely proportional to the thickness of the element. Accordingly, microfabrication allows eTECs to have high heat flux capability. Unlike chip-level cooling, it is possible for each functional block on an IC to be cooled independently to its desirable temperature setting. Electronic characteristics of various functional blocks and a simple model that captures characteristics of eTECs are discussed in this section.

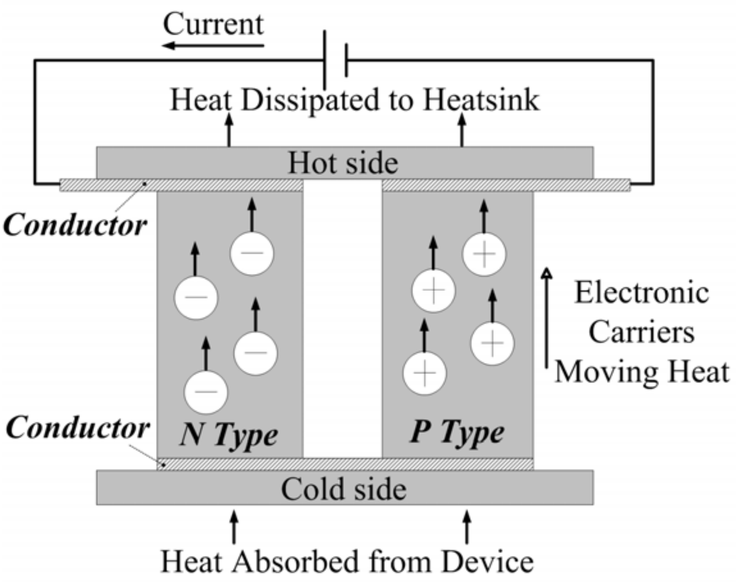


Figure 2-5: Schematic of a thermoelectric cooler.

2.2.1 Electronic Profile of Modern Microprocessors

Applying the simple transistor level model to a processor can be challenging. Each functional block is comprised of a large number of logic elements from thousands to millions of logic gates. The amount of power dissipated is an aggregate of the power of each element. Furthermore, the percentage of electronic power due to its leakage and active components depend on the functionality of digital blocks. The functionality impacts this ratio through the average activity factor, α. For instance, logic blocks such as an instruction decoder may have considerably different α as compared to memory blocks. Due to high-levels of logic activity under smaller area, logic blocks tend to have high power density with small percentage of leakage power. On the other hand, memory blocks attribute a large portion of its power to leakage. In this dissertation, McPAT, an integrated power, area, and timing modeling framework simulator, is used to explore to verify these electric characteristics. This simulator has shown results that modeled numbers track well with the published data for microprocessors [40].

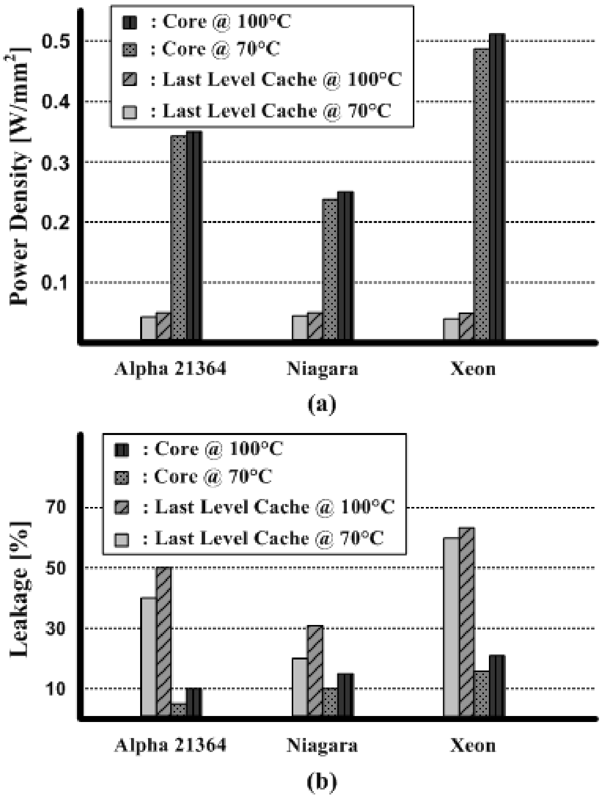


Figure 2-6: Electronic parameters of three different types of processors with (a) power density of core and last-level cache block at 70°C and 100°C and (b) leakage power ratio of core and last-level cache block at 70°C and 100°C.

Figure 2.6 illustrates power density and leakage percentage of a microprocessor's execution core and last level cache units for the 180nm Alpha 21364 processor [41], the 90nm Niagara processor [42], and the 65nm Xeon processor [43]. Clearly, core blocks have higher power density that is approximately 4 to 10 times higher than the last level cache blocks. Due to lower level of activity, last level cache dissipates a significant portion of its power to leakage with at least 30% and up to 60% of leakage at 100°C. The junction temperature of each block is determined by the total power density and the ambient temperature through a given thermal resistance. The figure shows the results at two different temperatures to account for the possibility of different thermal resistances.

The floorplan of the Xeon processor is shown in Figure 2.7 where the processor has two cores and a 16-MB shared L3 cache. Each core has a unified 1-MB L2 cache [43]. Note that approximately 70% of the total area is consumed by L-2 cache and L-3 cache. A more fine-grained breakdown of power density and leakage percentage of major functional components in the core block of the Xeon processor are shown in Figure 2.8. As expected, different functional blocks operate with different leakage and power density. Similarly, in general, blocks with high power density tend to have low leakage power percentage. The analyses in Chapter 4 use the data in Figure 2.6 and 2.8 to explore the impact of localized cooling on functional blocks.

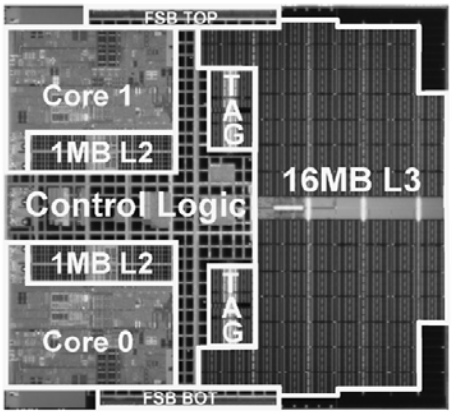


Figure 2-7: Floorplan of the Xeon processor. (Adapted from [43]).

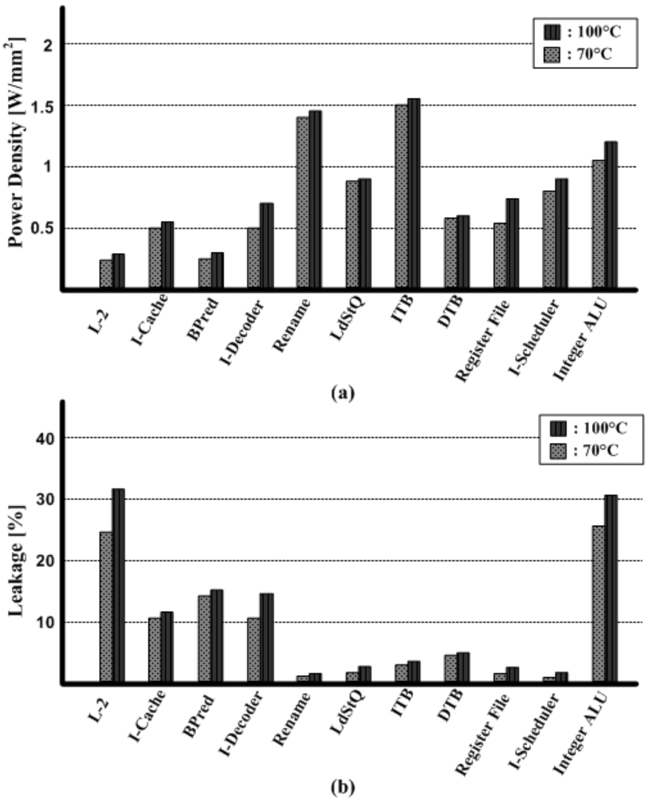


Figure 2-8: Electronic parameters of major function components in the core block of the Xeon processor with (a) power density at 70°C and 100°C and (b) leakage power ratio at 70°C and 100°C.

2.2.2 Embedded Thermoelectric Coolers

A configuration of microelectronic system with localized cooling is composed of a silicon wafer, a thermal interface material (TIM) with eTECs, a heat spreader, and a heatsink as shown in Figure 2.9. Conceptually, it is possible for the eTEC to be applied with fine granularity for each functional block of a processor. More realistically, the eTEC can be applied across coarse regions on an IC because of the substantial design overhead associated with using a large number of independent eTECs such as increased manufacture/package cost, additional power pins for cooling power, etc. This research considers both scenarios to explore whether substantial benefits can be garnered through a fine-grained temperature control.

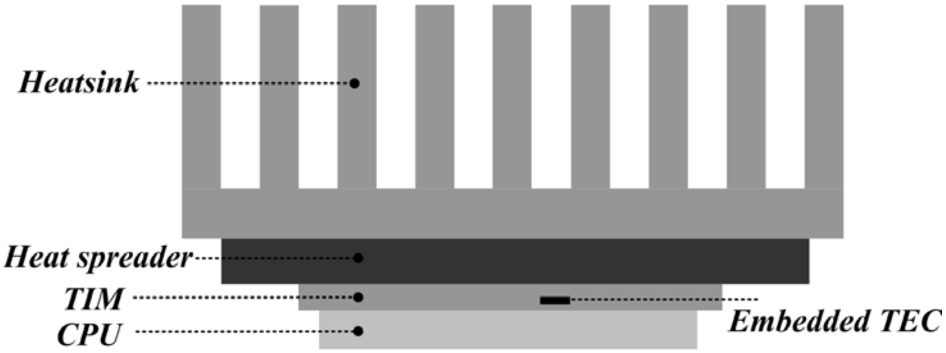


Figure 2-9: View of a microelectronic system with embedded TEC locally cooling a functional block.

Thermoelectric materials are characterized by their figure of merit ZT, defined by

$$ZT=\frac{S^2T}{RK} \quad (2.7)$$

where

S = Seebeck constant

R = electrical resistivity

K = thermal conductivity

T = temperature

Thermoelectric material needs to obtain low thermal conductivity in order to prevent heat losses through heat conduction between the hot and cold side. The material also needs high electrical conductivity to minimize Joule heating. A TEC absorbs dissipated electric power ($P_{electric}$) that is compensated by the heat conduction and Joule heating losses. Each of these components are shown as

$$P_{electric}=ST_cI-K\Delta T-\frac{1}{2}I^2R \qquad (2.8)$$

where T_c is the cold side temperature of the device, I is the electric current flowing through the device, and ΔT is the temperature difference between the cold and hot side. The input power of a TEC device is equal to

$$P_{cooling}=SI\Delta T+I^2R \qquad (2.9)$$

These parameters that characterize the performance of the eTEC are merged into a single model with the electronic power dissipation of a microprocessor as shown in Section 2.2.1. This single model considers the interdependence between the electronic and eTEC systems. With this model, we evaluate the impact of localized cooling on the power dissipation of functional blocks.

2.2.3 Performance of eTEC

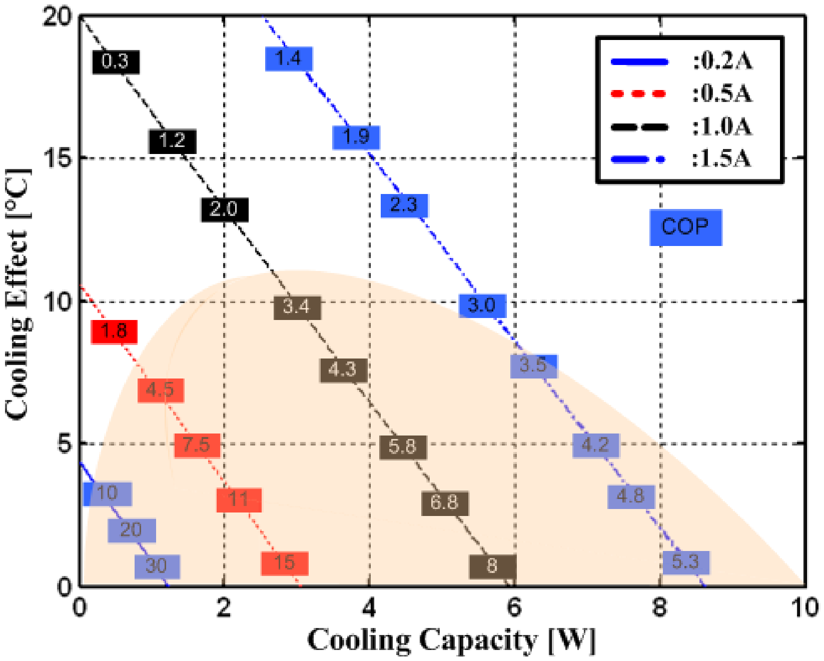


Figure 2-10: The performance curves of an embedded thermoelectric cooler with height and area of 100um and 6.25mm^2, respectively. (Adapted from[24])

As described in Section 2.1.2, refrigeration efficiency of any refrigeration cycle is measured by COP, defined by the ratio of the heat removed to the required power to remove the heat. In thermoelectric coolers, COP depends on operating conditions of I and ΔT as well as device parameters of K, R, and S as seen by the following equation.

$$\mathrm{COP}=\frac{P_{\mathrm{electric}}}{P_{\mathrm{cooling}}}=\frac{ST_cI\text{-}K\Delta T\text{-}\frac{1}{2}I^2R}{SI\Delta T+I^2R} \qquad (2.10)$$

We use parameters (Seebeck coefficient, electrical resistivity, and thermal conductivity) provided by [24], which has effective ZT of approximately 0.5. By adjusting the amount of current flow, the COP for different cooling capacity and associated cooling effect is obtained as shown in Figure 2.10 where cooling effect represents the amount of temperature difference obtained with cooling. The shaded area in the figure represents the range in which a high COP can be obtained using this particular device. Increase in input power allows the cooler to have a higher cooling capacity and cooling effect at the expense of efficiency. It is worthwhile to note that Figure 2.10 provides performance curves of the eTEC with the footprint area of 6.25mm^2 and the length of 100um. This dissertation investigates the effect of localized cooling on various types of blocks with different areas. Hence, we adjust the model of the eTEC depending on the size of the block under investigation. We also extrapolate from this data to model a hypothetical eTEC with higher performance that have effective ZT of 1.0 and 2.0 to explore the potential improvement through a better eTEC. The effectiveness of the thermoelectric modules can be improved by lowering electrical contact resistance and thermal contact conductance [23] - [25].

Using this model, Chapter 4 investigates whether localized spot cooling can lead to an overall system power saving when the cooling cost of eTECs is factored in. As the model includes the power dissipation for each block within a processor, we also investigate how much of an effect cooling has upon different types of blocks in terms of overall system power and temperature. This model serves as a useful method to evaluate the performance requirements of eTECs for electronic cooling.

2.3 Common Power Saving Techniques in Multi-Core Multi-Processor Systems

Minimizing power consumption of integrated circuits has been a hot topic in recent years. Along with reducing power consumption for mobile devices, significant amount of research has been directed towards power conservation for clusters of large-scale multi-server systems due to the tremendous growth in the number, sizes, and uses of data centers.

Using our system model from Section 2.1 and a highly efficient miniature-scale refrigeration system for electronic cooling, we illustrate the potential of power optimization of multi-core multi-processor systems based on different workload scheduling methodologies and power saving techniques. Results presented in this dissertation suggest that there exists an optimal methodology under the actively cooled environment that maximizes power efficiency while minimizing performance degradation. Furthermore, we combine our proposed methodology with the G/G/m-models to reduce the total power while meeting target SLA requirements by judicious capacity planning. This section reviews the basic power saving techniques in multi-core multi-processor systems and the G/G/m model from queueing theory.

2.3.1 Basic Power Saving Techniques

Powering-off cores that are not being used is the simplest way to reduce power consumption in multi-core processors. Stopping can be achieved in two ways. Most aggressive option is to save the architectural state and to completely power off the processor. This method is often known as power gating. Power gating the processor dissipates no power at all, but using this type of technique to power on/off comes at a price of long response time degradation since powering up a processor that is completely shut down requires up to 1000 cycles [44].

On the other hand, a simpler way to stop a core with minimal response time degradation of few clock cycles is to clock gate the core. This power saving technique stops active power dissipation, but since power is not entirely cut-off, the core continues dissipating leakage power. Advantages and disadvantages of using these two techniques at reduced temperature are explored in Chapter 6. Other special hardware supports for managing power in HPC processors exist, such as dynamic voltage frequency scaling, or thread migration, but this dissertation primarily focuses on the core stopping techniques.

Along with hardware-based power management techniques, energy usage of servers in the data centers can also be controlled through model-based energy-aware workload scheduling algorithms that distribute 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 [45] - [46]. 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 [47] - [48]. 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 operate at a certain utilization level to meet certain requirements often relies on a good model that successfully plans capacity depending on the upcoming workloads. For instance, *__G__eneral Arrival Process* / *__G__eneral Service Process* / *__M__ultiple Server* (G/G/m) model from queueing theory have been used to obtain useful measures to support capacity and workload planning of multi-processor systems in order to satisfy target Service Level Agreement requirements [49].

2.3.2 G/G/m Model for Capacity Planning

Execution Velocity and *Normalized Wait Time* are often used to obtain measures to support capacity planning of multi-processor systems. Figure 2.11 can be used to illustrate a generic G/G/m model for multi-server system, where each workload enters the queue of m-server system at an arrival rate of λ and leaves the system at a service rate of μ.

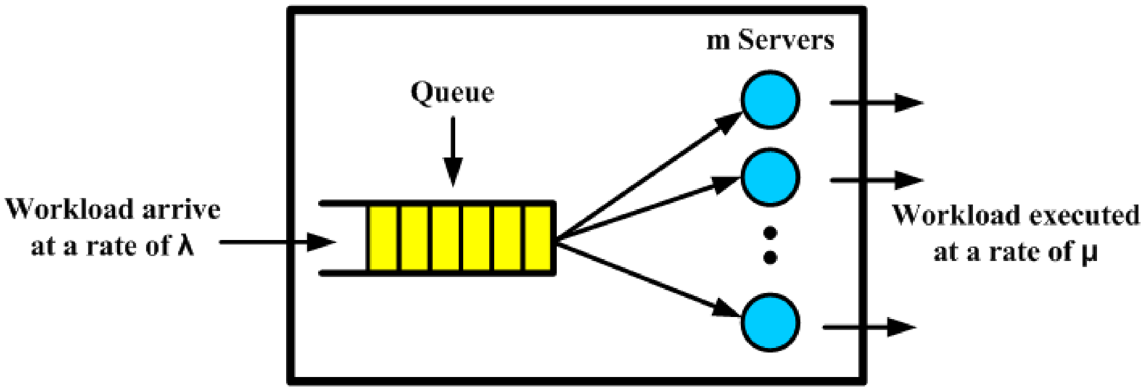

m Servers
Queue
Workload arrive at a rate of λ
Workload executed at a rate of μ

Figure 2-11: A Generic G/G/m model for multi-server systems.

The term *Execution Velocity* (E[V]) is used to describe the average ratio for the total amount of workload units that are served without any delay and is defined by Equation 2.11. The value of E[V] 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.

$$E[V] = 100 \times \left(1 - P\left[W > 0\right]\right) \quad (2.11)$$

The waiting probability, P[W>0], represents the probability that workload has to wait in the system from the delayed task. P[W>0] in G/G/m model can be approximately calculated by Equation 2.12.

$$P[W>0]=\frac{(m\rho)^{2}}{m!(1-\rho)}\cdot P_{0} \quad (2.12)$$

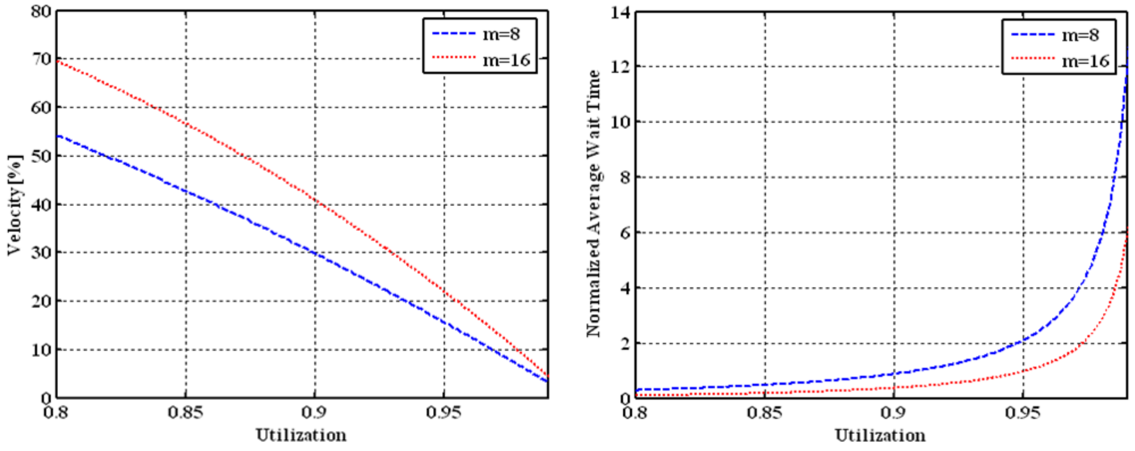
, where m is the total number of servers, ρ is the utilization ratio defined as λ/mμ, and P_0 is the probability for the empty G/G/m system defined as

$$P_{0}=\left[\sum_{k=0}^{m-1}\frac{(m\rho)^{k}}{k!}+\frac{(m\rho)^{m}}{m!}\frac{1}{1-\rho}\right]^{-1} \quad (2.13)$$

Next, we consider *Normalized Average Wait Time* (E[W]) where the term is used to describe the average wait time in the queue of the system normalized with respect to the service time of unit length of one unit. Several approximation formulas exist, but here we use the Allen/Cunneen formula [49].

$$E[W]=\frac{P[W>0]}{1-\rho}\cdot\frac{C_{A}^{\;2}+C_{B}^{\;2}}{2m} \quad (2.14)$$

where $C_A^2+C_B^2$ represents workload variability



(a) (b)

Figure 2-12: (a) Execution velocity and (b) normalized average wait time at different utilization for m=8 and m=16 processors.

As an example, Figure 2.12.a and 2.12.b show the results of execution velocity and wait time for 8 and 16-server system at different utilization. Notice how the number of the processors impacts execution velocity and wait time. Increasing the number of processors allows the system to operate at a superior performance at the expense of extra power dissipation. Underutilization of the system can also improve the performance. This results in a trade-off between high utilization of the processors (energy-conscious provisioning) and the target SLA requirements. Detailed analysis of minimizing overall power consumption of multi-core multi-processor systems while meeting target SLA requirements is presented in Chapter 6.

2.4 Summary

As the amount of power density continuously increases with the pace of technology scaling, the use of cooling through active heat removal is one potential option to combat both the power and thermal problems in HPC processors. Such cooling allows electronic systems to operate at a lower junction temperatures below the common forced-air cooling which in turn can alleviate thermally induced reliability issues and also reduce power consumption. A realistic system-level model that includes both the electronic characteristics and realistic performance of refrigeration systems is discussed in this chapter. The model is used to quantify the overall power consumption of a system that includes both the electronic and cooling power. This chapter focused primarily on two types of active cooling systems for electronic cooling: eTECs and vapor compression refrigeration systems. The two different cooling technologies differs in how they are applied in a system. Using eTECs have the benefit of full integration and provide flexibility of spot cooling different functional blocks within the processor. On the other hand, vapor compression refrigeration systems is more efficient than TECs. Physical size of this type of refrigerated system could be a limiting factor for certain applications, but refrigerated systems that fit within the small space of electronic chassis have been demonstrated [21] - [22]. Furthermore, this type of cooling system can often be extended to directly cool multiple dice in multiple chassis depending on the size of the compressor.

The model from Section 2.1 is extended to capture different relations and parameters of multi-core processor and the refrigeration system. Combining different power saving techniques, the model can potentially be used for minimizing power of multi-core multi-processor systems and investigate different workload scheduling methodologies. We also combine this model with the G/G/m-model to reduce both total power and response time degradation while meeting target SLA requirements. The overall simulator framework is shown in Figure 2.13. Our simulator provides overall system power consumption at different workload scheduling schemes for a given SLA requirements.

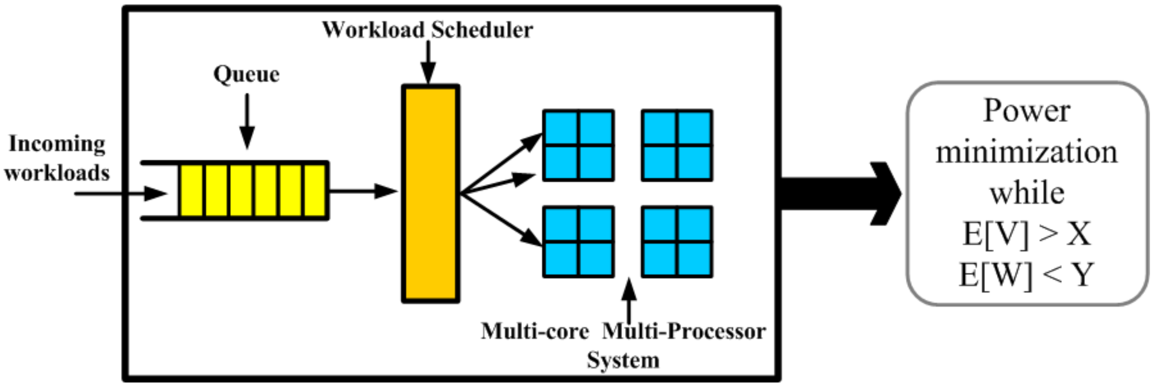


Figure 2-13: Overall simulator framework.

The next chapter explores the amount of obtainable power savings, optimal operating temperatures, and sensitivity to parameters of the electronic and refrigeration systems. Such analysis determines the feasibility of using refrigeration systems for electronic cooling.

CHAPTER 3

Analysis of Refrigeration Requirements for Electronic Cooling

Benefits of reduced temperature operation of CMOS circuits have been demonstrated in [50] - [51]. Lowered system temperatures have been extensively used for improving maximum operational frequency of high performance system but without optimizing for power. 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. The impact of cooling on nanometer-scale electronic systems has been recently discussed in [53] but uses a simple cooling model. This chapter describes analysis that indicates the feasibility of using refrigeration systems for electronic cooling using our compact model as described in Section 2.1.

In this chapter, various parameters that affect the overall power performance are discussed. Section 3.1 describes results from the model in a case study. Interestingly, the results from the analysis indicate an optimal operating temperature lie in the range of domestic refrigeration. Dependence on the characteristics of the electronics is presented in Section 3.2. From these results, performance requirements for the refrigeration are discussed for different electronic power-dissipation characteristics. Furthermore, sensitivity to ambient temperature, refrigeration efficiency, and thermal resistance of the package is presented in Sections 3.3. Finally, Section 3.4 summarizes the important findings and concludes with a brief discussion of the potential of refrigeration-assisted systems. The results from this chapter enable a designer to best choose or design the cooling system for the application.

3.1 Cooling a Microprocessor

This section uses the model in an example using a microprocessor as the heat load. The microprocessor operates at 3 GHz and dissipates 100W of power ($P_{electric0}$) using a 1.2V power supply at 80°C ($T_{junction0}$) without forced air cooling, which is summarized in Table 3.1. As in the microprocessor example in Table 2.2(b), $P_{leakage}/P_{electric}$ is assumed to be 50% at this operating point. The local ambient temperature ($T_{ambient}$) and the thermal characterization parameter of R_{jc} are 27°C and 0.2°C/W, respectively. Thermal resistance, R_{ca}, depends on whether it is air cooled or not. The refrigeration and fan performances and the relationship between thermal resistance and fan power are based on the data described in Section 2.1. Finally, the device parameters required for our simulator are extracted from BSIM model based on a 90nm CMOS technology process.

Parameter	
Initial $T_{junction0}$ [°C]	80
V_{dd0} [V]	1.2
Operating Frequency [GHz]	3
Leakage Percentage [%]	50
$P_{electric0}$ [W]	100
R_{jc} [°C/W]	0.2

Table 3.1: The model parameters used in the example of Section 3.1.

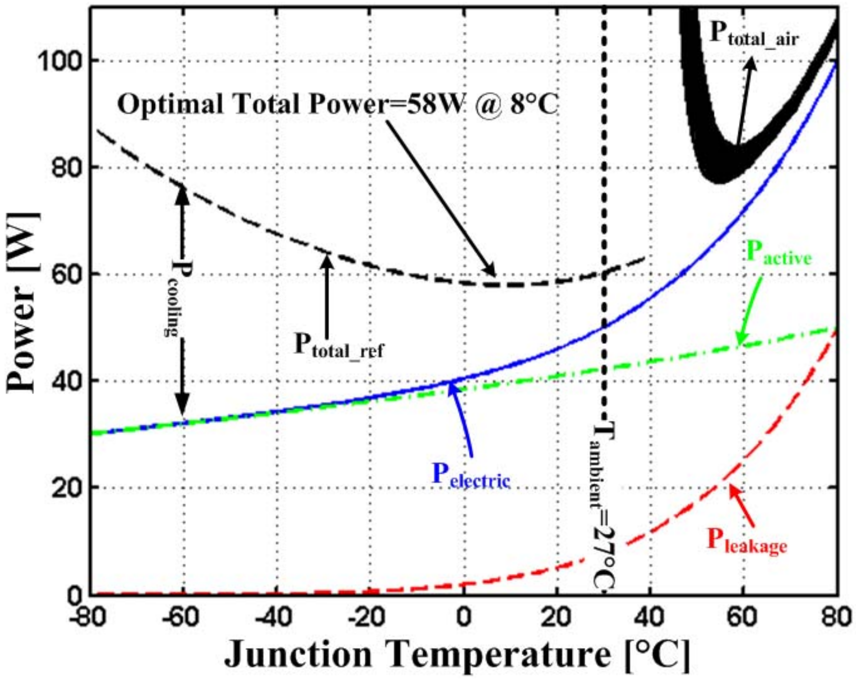


Figure 3-1: Optimization results of the system with forced-air and refrigerated cooling at Pleakage/Pelectric=50% @ 80˚C.

Total power is analyzed while sweeping the junction temperature and keeping the operating frequency constant. The result of the analysis is shown in Figure 3.1. The performance of forced-air cooling is plotted for comparison with refrigerated cooling. The analysis uses the assumption that the junction temperature is 80^{o}C. The performance curve for forced-air cooling is shown as a family of curves to account for the various possible fan size and R_{ca}. Decrease in R_{ca} at the cost of P_{fan1} shifts the curve to lower power. However, the amount of power reduction is limited by the obtainable R_{ca}. The optimal forced-air operating point for this particular heat load is roughly 77W at 55^{o}C.

Forced-air cooling is limited in the achievable $T_{junction0}$. As shown in the figure, refrigerated-cooling allows $T_{junction}$ to reach $T_{ambient}$ and below. Due to the increase in cooling cost at lower temperatures, the minimum power of 58W is obtained at 8°C. This performance is 45% better than the reference design and 25% better than the best-case forced-air cooled design. At this operating temperature, the active power is reduced by reducing the V_{dd} 1.1V. The most significant effect from cooling is the exponential decrease in the leakage power shown in Figure 3.1.

Analysis of the thermal circuit shows that T_{cold} of the refrigeration is 0°C at the optimal $T_{junction}$. At this temperature, the refrigeration efficiency is 48% of Carnot efficiency or COP of 3.33. If efficiency can exceed the bounds of the fitted curve in Figure 2.4, power savings can improve further. Similarly, the ratio of different workload ($P_{electric0}$) and $P_{leakage}/P_{electric}$ can impact the optimal temperature and efficiency substantially. These two scenarios are explored in the next two sections.

3.2 Varying $P_{leakage}/P_{electric}$ Ratios

In this section, the effectiveness of cooling at different $P_{electric0}$ and $P_{leakage}/P_{electric}$ ratios is analyzed. Two different analyses are performed. The first analysis maintains the $P_{electric0}$ while varying $P_{leakage}/P_{electric}$ ratio and the second analysis both varies the $P_{electric0}$ and the $P_{leakage}/P_{electric}$ ratio. Other parameters associated with the analyses are same as in Table 3.1.

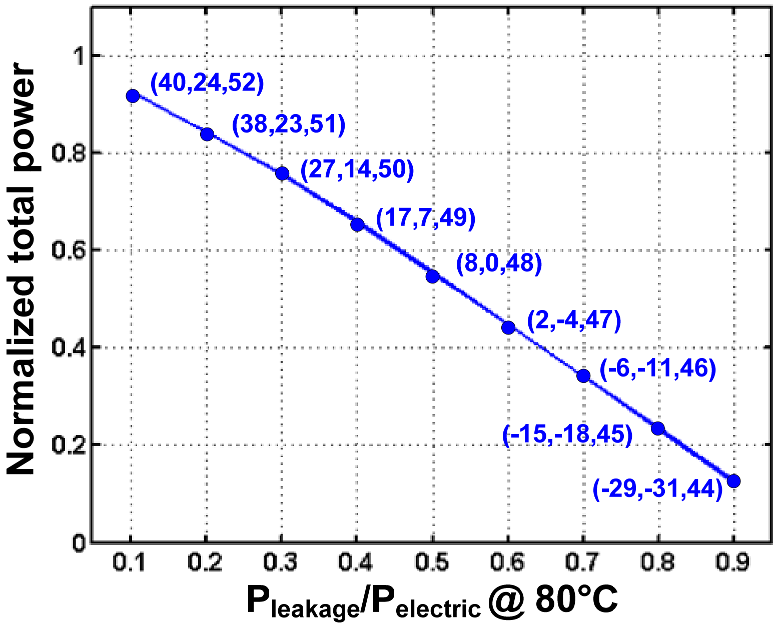


Figure 3-2: The dependence of the $P_{leakage}/P_{electric}$ ratio on the optimal total power based on the simulation using the model. (x,y,z) correspond to optimal operating points of $T_{junction}$ [°C], T_{cold} [°C] , and the percent of Carnot [%], respectively.

In the first scenario, the total $P_{electric0}$ is maintained at 100W with $T_{junction}$ at 80˚C with forced air cooling to correspond to the current practice of constraining the maximum power of the digital processor. The results of the analysis are shown Figure 3.2. Each data point indicates the power normalized to 100W. The figure includes the cold-end temperature, the junction temperature and the efficiency of the refrigeration system for each point. The power savings is roughly proportional to the percent of leakage. Figure 3.1 helps explains the nearly linear relationship. Because of the exponential relationship between

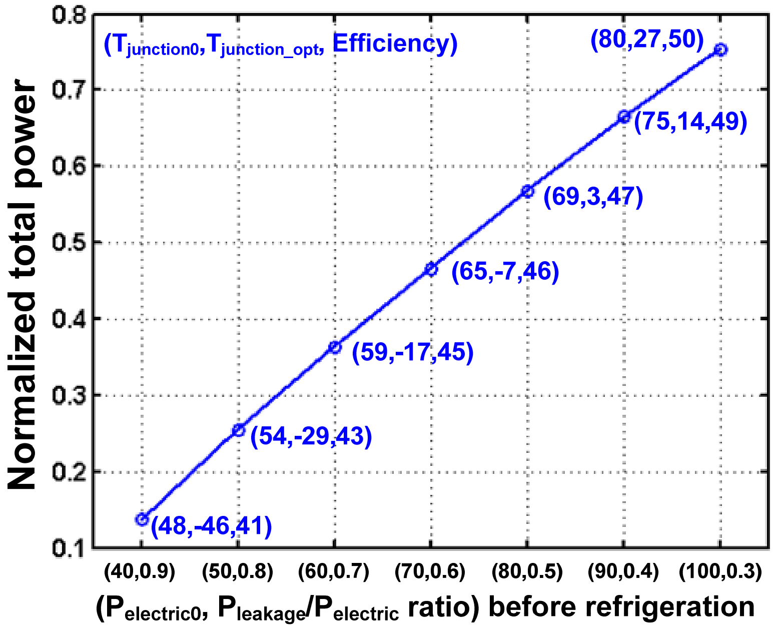


Figure 3-3: The dependence of the $P_{electric0}$ and leakage ratio on the optimal total power based on the simulation using the model. (x,y,z) correspond to junction temperature before refrigeration $T_{junction0}$ [°C], optimal operating points of $T_{junction}$ [°C] , and the percent of Carnot [%], respectively.

leakage power and temperature, most of the power saved at the optimal temperature is due to the leakage power. The analysis indicates only a modest power improvement when the system power is entire due to active power. The power savings due to active power is essentially offset by the cooling power. Note that the optimal operating temperature is lower for the system with higher leakage power.

Second scenario investigates varying $P_{electric0}$ and the $P_{leakage}/P_{electric}$ ratio to simulate realistic application workload. A processor runs different applications or instructions and the instantaneous activity factor changes. In this analysis, $P_{electric0}$ varies from 40W to 100W. The $P_{leakage}/P_{electric}$ ratio is chosen to be 90% when $P_{electric0}$ is at 40W corresponding to the system in standby mode. The ratio decreases as $P_{electric0}$ increases reaching down to 30% when $P_{electric0}$ is 100W. The results of the analysis are shown Figure 3.3. Each data point indicates the power normalized to the power before refrigeration. The figure includes the junction temperature before refrigeration, optimal operating points of $T_{junction}$ and the

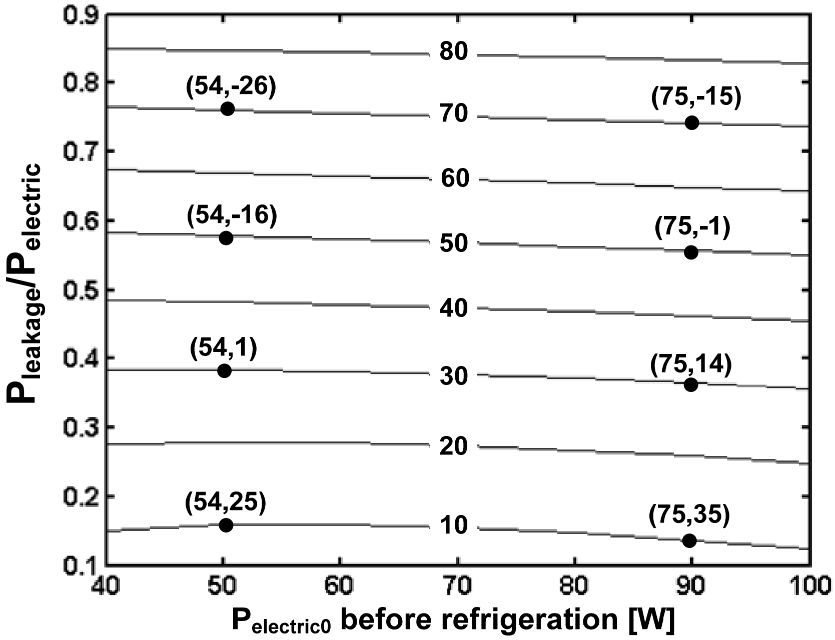
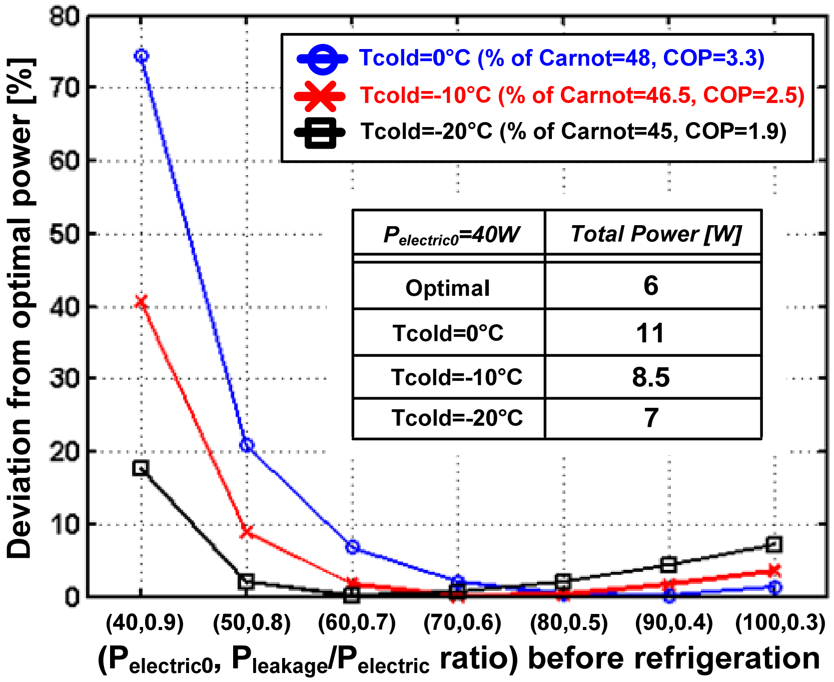


Figure 3-4: Contours of the percentage of power reduction for different $P_{leakage}/P_{electric}$ and $P_{electric0}$ before any refrigeration. (x,y) correspond to $T_{junction0}$ [°C] before refrigeration, and optimal junction temperature after refrigeration $T_{junction}$ [°C], respectively.

efficiency of the refrigeration system. The junction temperature before refrigeration, $T_{junction0}$, changes because of the lower heat load. It is important to observe that the curve exhibits the same proportional relationship as in Figure 3.2.

A different view of this analysis is shown in Figure 3.4. At each workload ($P_{electric0}$), the power savings is calculated for different $P_{leakage}/P_{electric}$ ratios. The figure shows contours of constant power reduction. Note that the optimal temperature has a slight dependence on the value of $P_{electic0}$. When $P_{electric0}$ is small, the optimal temperature is lower due to the lower $T_{junction0}$. Lastly, in all of the above analyses, when $P_{leakage}/P_{electric}$ ratio is between 30 and 70%, the optimal operating temperatures are in the vicinity of domestic refrigeration systems (approximately 0^{o}C ± 20°C).

So far, the analysis optimizes the temperature for different $P_{leakage}/P_{electric}$ ratios assuming the highest achievable COP for that temperature from the fitted result in Figure 2.4. Below, the analysis is repeated using a single refrigeration system. Two implementation approaches are examined. First, a constant cold-end temperature is maintained regardless of the workload. The temperature is chosen to be at the peak COP of a refrigeration system. Second, the operating temperature is adapted while using the realistic COP of a single refrigeration system.



$P_{electric0}$=40W	*Total Power [W]*
Optimal	6
Tcold=0°C	11
Tcold=-10°C	8.5
Tcold=-20°C	7

Figure 3-5: Effect on the overall system performance when T_{cold} is kept to a fixed temperature. Varying fixed temperatures are used. The COP at each T_{cold} corresponds to the achievable COP of that temperature. The table shows the total power after refrigeration when $P_{electric0}$ is 40W.

In the first approach, T_{cold} is kept at a fixed temperature that is within the range of domestic refrigeration compartments (-20°C ~ 0°C). As the $P_{leakage}/P_{electric}$ ratio is varied, Figure 3.5 shows the amount of deviation from the previous optimal results of Figure 3.3. Note that when T_{cold} is fixed at 0°C, the amount of deviation increases sharply for $P_{electric0}$ less than 70W but stays close to the optimal value when $P_{electric0}$ is greater than 70W. This behavior is because the optimal cold-end temperature decreases substantially as the $P_{electric0}$

decreases. As shown in the figure, a better choice of T_{cold} is -20°C which results in power that are within 20% of the optimal point. The choice of an optimal fixed T_{cold} depends on the system's range of $P_{electric0}$ and $P_{leakage}/P_{electric}$ ratios during operation (i.e. the range of the average activity of a processor). The design of the refrigeration systems should target this temperature. This analysis also serves as an indication of the sensitivity of the optimal operating point. The flatness of the optimum shows that small deviations from the optimal temperature do not lead to substantial degradation in performance. The table embedded in the figure also shows the total power when $P_{electric0}$ is 40W to show that the large percentage deviation from optimal is due to the small total power after refrigeration.

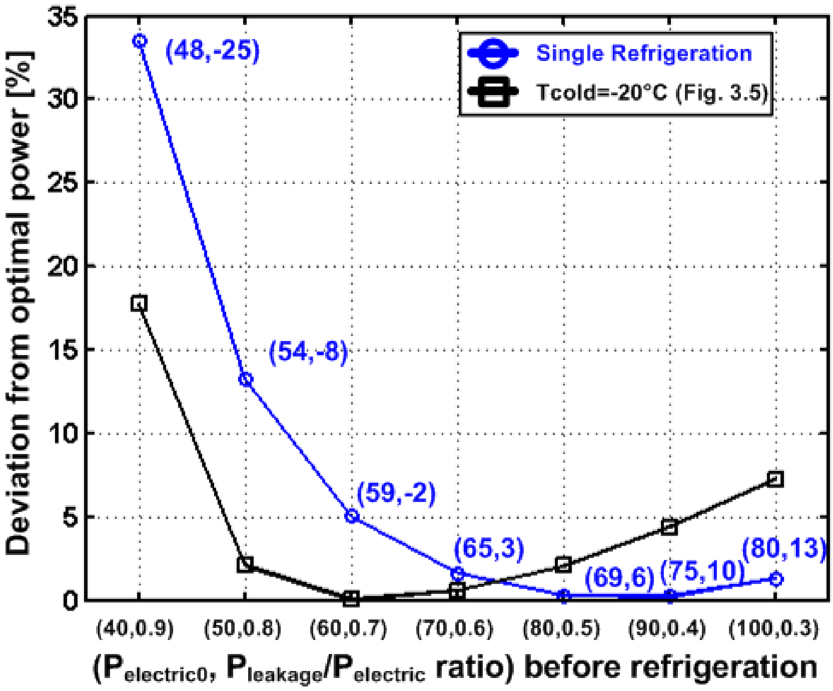


Figure 3-6: The effect on the overall system performance, assuming single refrigeration system is being used. (x,y) correspond to junction temperature before refrigeration $T_{junction0}$ [°C], adapted $T_{junction}$ [°C] using single refrigeration system.

The second approach adapts T_{cold} using the realistic COP data of a single refrigeration system for different $P_{leakage}/P_{electric}$ ratios. The analysis is performed using the refrigeration efficiency data of [36]. The refrigeration system has been optimized for domestic refrigeration temperature of 0°C but the COP falls off at other temperatures. Figure 3.6 shows that total power is within 5% of the optimal result for $P_{electric0}$ greater than 60W. Interestingly, the adapted $T_{junction}$ when $P_{electric0}$ greater than 50W ranges from -8°C to 13°C and deviates only slightly from the nominal design point of the refrigerator (0°C).

Therefore, adapting T_{cold} does not lead to substantial improvements as compared to maintaining a constant T_{cold} for a large range of $P_{leakage}/P_{electric}$ ratios. Unless the refrigeration system has a broad range with high COP near the targeted temperature, the benefits of adapting the optimal temperature may not be worth the design overhead. Next section explores sensitivity of cooling parameters and temperatures on the overall performance.

3.3 Varying Cooling Parameters

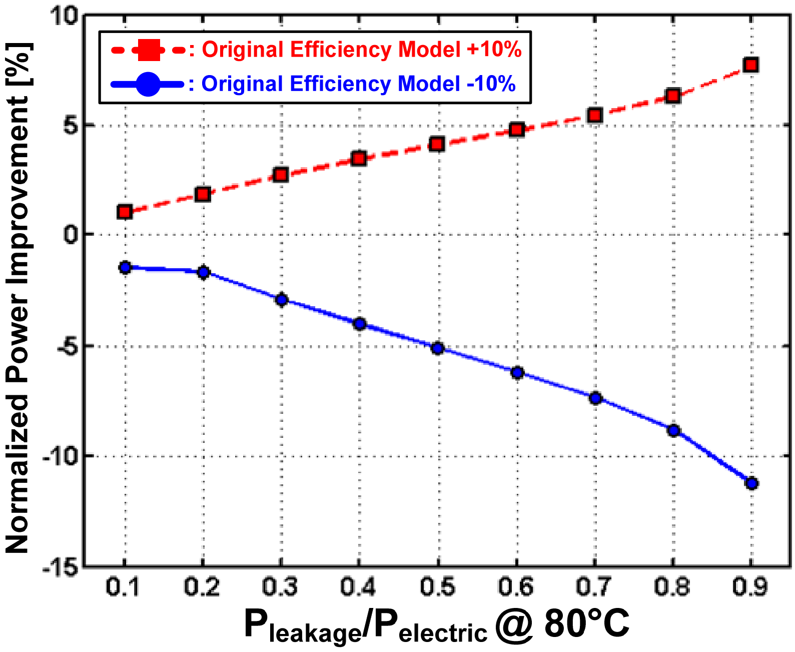


Figure 3-7: The sensitivity of the power to a change in refrigeration efficiency data. The results are normalized with respect to Figure 3.2.

The sensitivity of the power due to different refrigeration efficiency data is investigated to show that the power can improve further if the efficiency can exceed the bounds of the fitted curve in Figure 2.4. In this analysis, the refrigeration efficiency data is assumed to be ± 10% from the bounds of the original curve. Figure 3.7 shows that the magnitude of power savings is sensitive to the $P_{leakage}/P_{electric}$ ratios. A reduction in COP has a substantial impact on the power savings. At low $P_{leakage}/P_{electric}$ ratios, a 10% reduction in efficiency (η) essentially leads to no improvement in power from refrigeration. This result indicates that less efficient cooling approaches may yield no power benefit unless the system has a large leakage component. As discussed in Chapter 4, a minimum efficiency of thermoelectric cooling is expected to be necessary to improve overall power consumption.

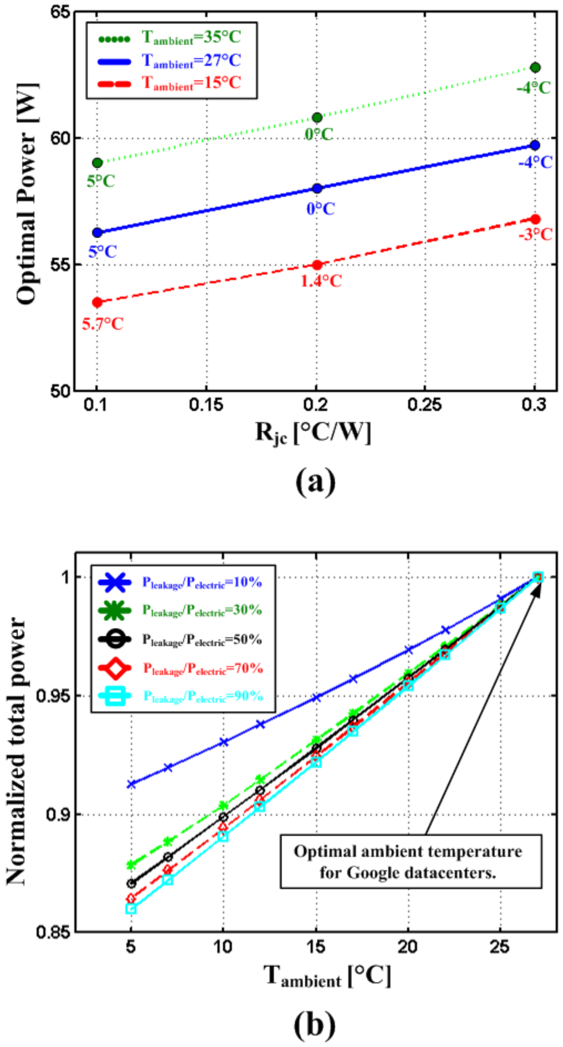


Figure 3-8: (a) System performance while varying R_{jc} and $T_{ambient}$ at $P_{electrio0}$ and $P_{leakage}/P_{electric}$=50%. The corresponding optimal T_{cold} are also shown. (b) The effect on the overall performance from varying $T_{ambient}$. Normalization is done with respect to the total power when $T_{ambient}$ is at 27°C.

The sensitivity to R_{jc} and $T_{ambient}$ are explored using the model. The study in [35] indicates that typical values of 0.1 ~ 0.3 °C/W for R_{jc} are obtainable with careful design choices, and thereby are used in our analysis. Smaller values of R_{jc} enable the heat to transfer more efficiently, enabling a higher T_{cold}. The results in Figure 1.8(a) indicate that improving the thermal resistance yields only a modest improvement on power savings (~5%).

A decrease in $T_{ambient}$ also results in additional power reduction. At lower $T_{ambient}$, T_{hot} decreases which consequently increases the refrigeration efficiency for a given T_{cold}. Simulation results (Figure 3.8(b)) indicate that power savings is roughly proportional to the ambient temperature. Similar to our previous analysis, however, the power consumption of the refrigeration system that cools the ambient must be taken into account in order to quantify the overall benefit to the system. Research for optimal temperature of large data centers [54] shows an optimal $T_{ambient}$ of 27°C indicating that any benefits are likely offset by the ambient refrigeration.

3.4 Summary

This chapter incorporates realistic refrigeration and electronic data in analyzing the impact of sub-ambient cooling on the total power of processor systems. This type of active cooling can not only lead to substantial power-performance improvement for the electronics, but the overall system performance improves. The power savings is roughly proportional to the ratio due to the sensitivity of leakage power to temperature. The optimal temperature primarily depends on the amount of leakage power and the ratio of leakage to total electronic power. Using realistic refrigeration parameters, our analysis found that the target temperature is near that of domestic refrigeration.

While adapting the temperature to the operation of the electronics can lead to better power, the difference is small as compared to using a constant temperature due to the poor refrigeration COP as the temperature deviates from the target refrigerator design. Refrigeration system needs to be designed with high COP across a broad range of temperature for the adaptation to be beneficial.

The results of the analysis also indicate that enhancements to refrigeration systems at domestic freezer temperatures can potentially lead to even greater power savings. Additionally, since the performance of refrigerated cooling depends on the characteristics of the electronics, different circuit and architectural techniques can be explored to better take advantage of the sub-ambient temperature.

Based on the analysis presented in this chapter, a miniature-scale refrigeration system for electronic cooling is developed and is discussed in Chapter 5. Before discussing measured results, the next chapter explores whether power improvement is possible by cooling individual blocks within a processor. Characteristics of embedded thermoelectric coolers as presented in Chapter 2 are incorporated into the analysis in the next chapter.

CHAPTER 4

Analysis of eTEC-based Cooling

Advancements in embedded thermoelectric coolers (eTECs) are enabling integration with IC processing for localized cooling. Recent interest is in using eTEC to locally and selectively cool an area with high power density to resolve thermal problems in integrated circuits. One study claim that localized cooling of an IC may result in reduced power consumption since lowering the temperature can reduce power consumption [24]. This chapter addresses whether localized spot cooling can also lead to an overall system power saving when the cooling cost of eTECs is factored in. Furthermore, the effectiveness may vary depending on different types of blocks.

This chapter explores the optimal operating temperatures and the amount of total power reduction achievable from localized cooling using the model in Section 2.2. Dependence of the operating point on parameters of the electronics and cooling systems is determined. As shown in this chapter, enhancements to thermoelectric elements are critical in order to obtain a reasonable amount of overall system power saving.

The chapter is organized as follows. The effect of localized cooling on various types of blocks using different thermoelectric elements is presented in Section 4.1. From these results, Section 4.2 discusses the amount of total power saving for current and future processors. Section 4.3 summarizes with a brief discussion of future potential of eTEC systems.

4.1 Cooling a Microprocessor Using eTECs

By using the performance curve from an eTEC with ZT=0.5 (Figure 2.10), the total power is analyzed while sweeping the junction temperature and keeping the operating frequency constant. Our first example considers the impact of cooling the *Branch Predictor* unit and the *L-3 Cache unit* of the Xeon processor from Section 2.2. The *Branch Predictor* dissipates 0.19W of power ($P_{electric0}$) from a 1.25V power supply while operating at 3.4 GHz. The power density of this particular block is 0.34W/mm^2. At this operating point, $P_{leakage}/P_{electric}$ is 16%. The power density and the $P_{leakage}/P_{electric}$ ratio of the *L-3 Cache* unit are 0.05W/mm^2 and 67%, respectively. Here, only a small portion of the *L-3 Cache* unit is cooled.

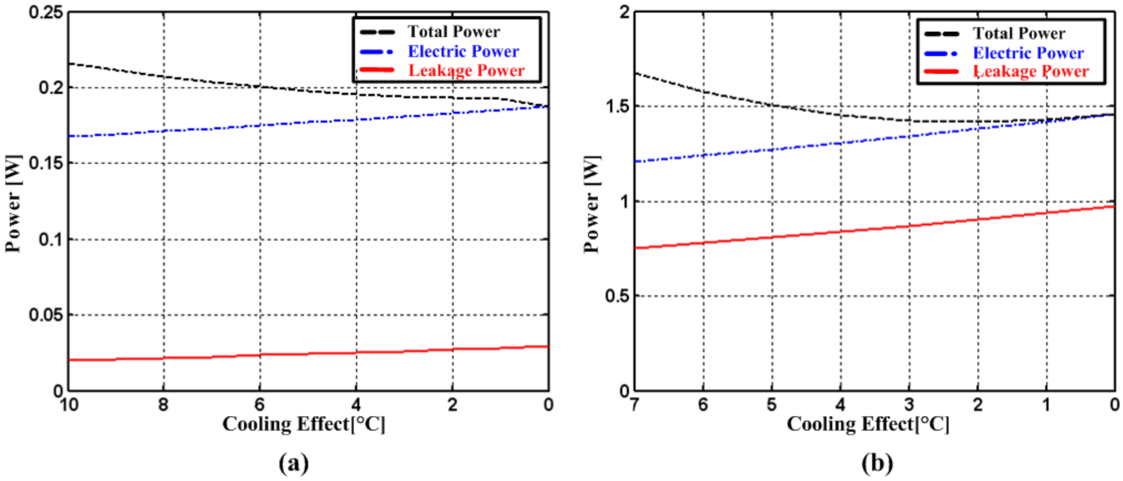


Figure 4-1: Optimization results of localized cooling (a) Branch Predictor and (b) L-3 Cache.

The result of the analysis is shown in Figure 4.1. Localized cooling allows $T_{junction}$ to operate at a lower temperature at the expense of cooling power. Total electric power decreases at lower temperature, but total power consumption increases when including the cooling cost. Due to the sharp increase in cooling cost at lower temperatures, cooling the *Branch Predictor* unit does not provide any system power benefits. On the other hand, *L-3 Cache* unit dissipates the minimum total power of 1.41W with 2°C of cooling effect. This performance is only marginally (3%) better than the reference design. The amount of power reduction is limited by the obtainable COP. This particular example indicates that for localized cooling to work in terms of overall power perspective, both high COP and high cooling effect eTECs are required. Otherwise, due to the increase in cooling power cost, localized cooling does not yield significant system power and thermal benefits.

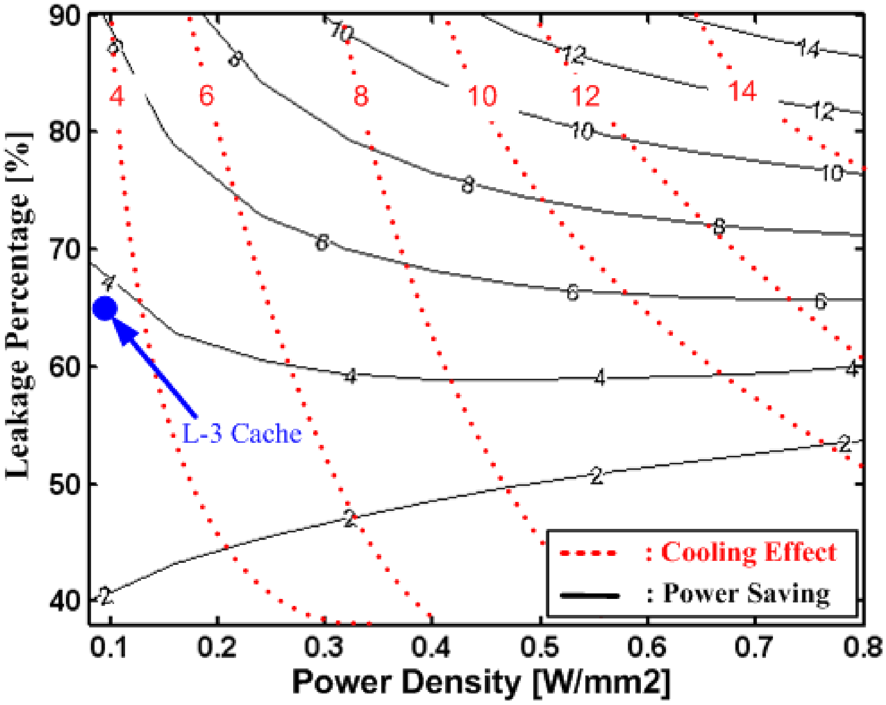


Figure 4-2: Contours of the percentage of power reduction and amount of cooling effect for different $P_{leakage}/P_{electric}$ and power before any refrigeration.

To account for different functional blocks, Figure 4.2 illustrates overall power performance for different $P_{leakage}/P_{electric}$ ratios and amounts of power density. The figure shows contours of the power savings [%] as well as the amount of cooling effect [°C] at various optimal power points. It can be shown from the figure that the amount of power savings and cooling effect depends on the characteristics of each functional block. Because of the exponential relationship between leakage power and temperature, the power saving of using TEC increases when the system has a large leakage component. Simulation shows that more noticeable power savings of 5-10% start to appear for systems with high leakage ratio. The amount of power saving is limited to 14% even for blocks entirely dominated by leakage. This analysis also indicates that the power saving has a dependence on power consumption. Note also that the amount of cooling effect increases with leakage percentage and power consumption but ranges only from 4 to 14 degrees at each optimal power point. Most functional blocks have leakage percentage that is <40% (see Figure 2.8). This implies that the use of this particular eTEC for cooling the processor core does not result in overall system power saving and may not be worth the design overhead.

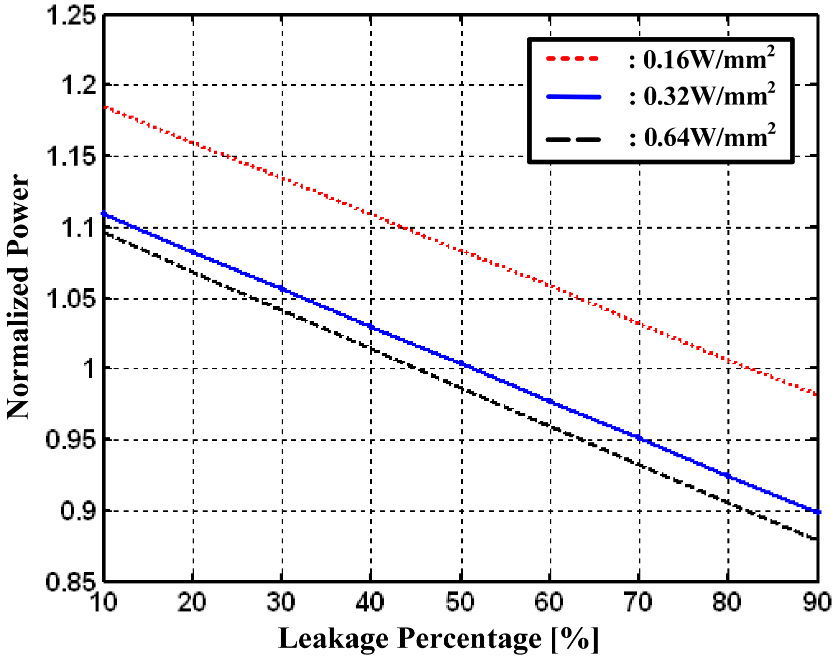


Figure 4-3: Effect on the overall system performance to obtain 10 degree of cooling effect using eTEC with thermoelectric material of ZT=0.5.

It is important to recognize that temperature reductions of 10°C can be important in cooling local hotspots to reduce the junction temperature for higher reliability. For this purposes, some amount of power penalty can be tolerated. We extend the analysis to consider the respective power penalty when obtaining 10 degrees of cooling effect for different $P_{leakage}/P_{electric}$ ratios and power. The simulation result is shown in Figure 4.3. We chose three possible hot spots with different amounts of power density: 0.16W/mm^2, 0.32W/mm^2, and 0.64W/mm^2. The results indicate that the block with large power density and high leakage power percentage gives the least amount of power penalty. In fact, in the case of leakage ratio of >45% with power consumption of 0.64W/mm^2, we see that no extra power is required to obtain 10 degree of cooling.

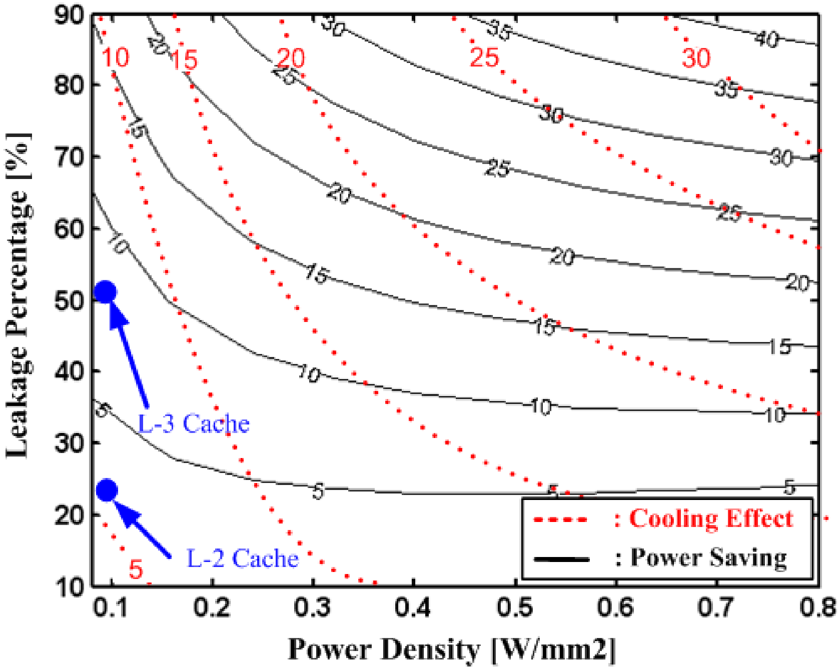


Figure 4-4: Contours of the percentage of power reduction and amount of cooling effect for different $P_{leakage}/P_{electric}$ and power before any refrigeration with ZT=1.

Since an eTEC with ZT=0.5 does not appear to demonstrate substantial power savings, we repeat the analysis using a hypothetical thermoelectric material with ZT=1. For a given COP, the effective cooling approximately improves by a factor of two by doubling the ZT. Other parameters associated with the analyses remain the same. Figure 1.4 shows the maximum power savings and their associated cooling effect for different $P_{leakage}/P_{electric}$ ratios and power density. When compared to the results in the previous section, power savings of at least 15% is realizable when the components have leakage ratio that are greater than 50% with power density of 0.5W/mm^2 or greater. This performance is 7X better than the performance obtained with thermoelectric element of ZT=0.5.

Moreover, the cooling effect at each optimal power point has increased substantially. In the case of components with high leakage and power, it is possible to lower the temperature by 10 without requiring a higher total power. In fact, for most of the different functional blocks in a microprocessor only a small amount of extra power is required to obtain 10 degrees of cooling. As apparent from these simulations, by using an eTEC with higher ZT, cooling can be used without power cost to improve reliability and localized cooling can actually result in reduced overall system power.

4.2 Results of Cooling a Microprocessor

The eTECs are assumed to be locally applied with fine granularity to each functional block in the processor for spot cooling. Each block is cooled to its optimal temperature setting. Our results of cooling the functional blocks from Figure 2.8 using ZT of 1 are summarized in Table 4.1. As expected, blocks with the larger portion of leakage power consumption have the strongest influence from localized cooling. For instance, cooling the

ZT=1	Power Saving [%]	Cooling Effect [°C]
L-3 Cache	7	6
L-2 Cache	2	8
I-Cache	-5	10
BRED	-2	10
I-Decoder	-3	10
Rename	-20	10
LdStQ	-11	10
ITB	-21	10
DTB	-8	10
Register File	-12	10
I-Scheduler	-12	10
Integer ALU	-8	10

Table 4.1: Simulation result of fine-grain cooling each of the functional blocks from Figure 2.8.

Level-2 Cache unit *Level-3 Cache* unit resulted in the maximum power saving of 2% and 7%, respectively. Associated cooling effects at their optimal operating points are 8 and 6°C. On the other hand, cooling other components does not provide any power benefits, and extra power would be needed to obtain temperature reductions. Table 4.1 shows the respective power penalty to obtain 10 degrees of cooling effect, represented by the negative percentage. For an example, cooling hotspots like the *Rename* unit would require extra 20% of power to reduce the junction temperature by 10 degrees. These results demonstrate that cooling with fine granularity is not necessary. Analyses provide a subtle conclusion that cooling the cache blocks result in the performance that is better than logic based blocks.

This observation counteracts the underlying assumption of using eTEC for cooling hot spots.

Compared to non-cooled processor, cooling only the cache units result in total power performance that is 2% better. Power savings from localized cooling is not as significant as from chip level cooling where chip level cooling using vapor compression refrigeration system would result in total power saving of >30%. Nevertheless, localized cooling the right element can give some worthwhile improvements while providing the benefit of full integration.

Tech Node (nm)	Level-2 Cache		Level-3 Cache	
	ZT=1	ZT=2	ZT=1	ZT=2
65	2	5	7	37
45	7	17	10	50
32	13	27	16	60
22	14	29	18	62

Table 4.2: Overall power saving [%] of cooling L-2 and L-3 cache at different technology nodes.

The processor designed in 65-nm technology is scaled down to 45nm, 32nm, and 22nm technology nodes by assuming the area is proportional to the square of the feature size using the device parameters from the ITRS [18]. Table 4.2 provides the amount of power saving for the *Level-2 Cache* unit *Level-3 Cache* unit using thermoelectric material with ZT of 1 and 2. Due to the increase in leakage percentage with technology scaling, amount of power saving increases. Using eTECs of ZT=1 for cooling the *Level 3 Cache* is projected to improve the amount of power saving with technology scaling, reaching up to 16% for 32nm technology and beyond. Table 4.2 also summarizes the result of using higher ZT of 2. As expected, significant increase in power savings can be obtained.

4.3 Summary

This chapter explores the benefits of using a different cooling mechanism, eTEC, to solve power and thermal problems when factoring in the cooling cost of eTECs in the total system power. The complete system-level model that includes both the realistic electronic characteristics of various functional blocks and the performance of eTECs from Section 2.3 is used. Analysis indicates that the performance primarily relies on the amount of leakage power and the ratio of leakage to total electronic power. Power savings using eTECs are not as significant as cooling using other means such as a vapor compression cycle at the chip level but they have the benefit of allowing a fully integrated solution. The analysis shows that with ZT=0.5, cooling the right element (the cache) can give a modest 3% improvement. Enhancing the performance of the eTEC provides a much more compelling performance improvement in terms of both power (>10%) and reliability. Finally, since the amount of leakage power increases with technology scaling, using eTECs may be a possible power/thermal solution for future electronics.

CHAPTER 5

Measured Results Using Refrigerated Cooling

We demonstrated in Chapter 3 that active cooling not only can lead to substantial power-performance improvement that includes the cost of cooling power. Moreover, there exist optimal operating temperatures. The power savings including the cost of cooling is roughly proportional to the leakage power ratio due to the sensitivity of leakage power to temperature.

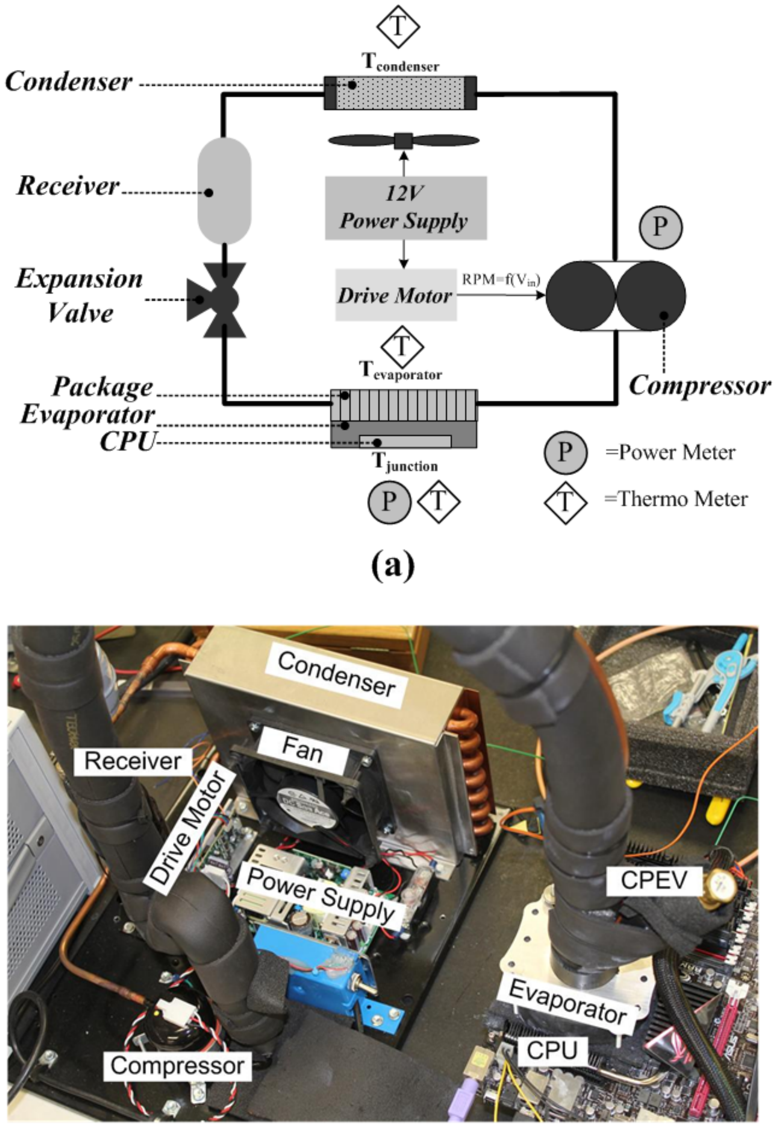
This chapter describes an experimental setup to validate the analysis in Chapter 3. Achieving power improvement requires a highly efficient dedicated refrigeration system for cooling the IC. Several notable research efforts have demonstrated miniature-scale vapor compression refrigeration system for electronic cooling that can fit into a small space of an electronic chassis. For example, one of these refrigeration systems [21] has shown an evaporator temperature range of 8 to 22 ºC while having the cooling capacity of 120 to 280W with the COP ranging from 3 to 4.5. The authors confirmed that a highly efficient compressor that is small enough to fit within a space of an electronic chassis can greatly enhance electronic systems.

Our work builds upon and expands this work by developing a miniature-scale refrigeration system that is capable of operating at a lower evaporator temperature with higher efficiency. Using this refrigeration system, this chapter focuses on developing a complete system that could characterize both the electronic and cooling performance. Instead of using a heat source which simulates integrated circuits, the refrigeration system is mounted onto a multi-core microprocessor for overall system performance characterization. This experimental setup allows us to explore and demonstrate the processor performance and power reduction at lower operating temperatures.

A processor that dissipates 175.4W of maximum power with 30% electronic leakage power operating at 97°C is cooled using our refrigeration system. Measurements show that with a minimum refrigeration COP of 2.7, the processor operates with junction temperature <40°C and offers a 25% total system power reduction over the non-refrigerated design. This experiment is the *first demonstration* of active cooling that lead to reduced total wall power. With an improved compressor that maintains the COP across a broad range of cooling capacity, our analysis shows that at least >13% of total power is saved across the entire range of processor utilization.

This chapter begins with a description of the overall system and our experimental setup for the overall system performance characterization. Based on the measurements and the system-level model, we discuss the amount of total power saving for processors. Finally. the chapter ends with suggestions for future electronic cooling systems.

5.1 Overall System Description



(a)

(b)

Figure 5-1: (a) Layout and (b) photograph of the refrigeration system for electronic cooling.

A layout and photograph of the refrigeration system for electronic cooling is shown in Figure 5.1. A configuration of the refrigeration system consists of a compressor, a condenser, a constant pressure expansion valve (cpev), a cold plate evaporator, and a cooling fan. A 12V power supply is used to provide the required power. Additionally, a motor drive board is installed to control the compressor speed, and a receiver is installed to guarantee that only refrigerant liquid enters the cpev. 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. The refrigeration system is configured to cool the microprocessor at different heat loads and temperatures.

A commercially available yet affordable miniature refrigeration compressor in the hundred-watt range that fits within the small space of an electronic chassis is used [55]. The dimension of the compressor is 5.58 cm in diameter and 7.75 cm high, and weighs only 590 g. The compressor is driven by a sensorless brushless DC motor running at 12V DC. The compressor is able to operate at an evaporator temperature range of -18 to 24°C and a condenser temperature up to 71°C. The compressor speed ranges from 1800 to 7000 RPM controlled by the drive motor with an external 20kΩ potentiometer.

The automatic cpev regulates the liquid refrigerant flow entering the evaporator at a constant outlet pressure at all load conditions, matching the compressor capacity. By adjusting the amount of flow using the knob of the cpev, the temperature of the evaporator can be controlled. Although, widely available microchannel condenser could have been installed, our prototype uses a condenser with dimensions 50mm X 230mm X 180mm that has heat rejection capacity of 260W at 90 CFM airflow. A 12V DC fan provides sufficient amount of air flow across the condenser. The fan rejects the hot air into the ambient environment. Finally, this cooling system is charged with 120g of R-134 refrigerant.

To guarantee a small difference between the heat input from the microprocessor and the cooling capacity, the interface between the cold plate evaporator heat exchanger and the heat source must be tightly sealed and insulated to reduce any exposure of the cold surfaces to the ambient environment. Different layers of insulation forms are placed between the board bracket and the back side of the motherboard and at the top surface of the motherboard around the microprocessor socket. The design avoids leaving any air pockets that might result in building up condensation. Extra layer of the insulation is placed over the nozzle of the evaporator copper head. Finally, the evaporator is mounted over the TIM pasted microprocessor with the mounting screws for a firm contact and pressure.

5.2 Performance of the Refrigeration System

The performance of the compressor dictates the overall refrigeration system performance. Refrigeration power which accounts for the Carnot cycle and compressor power is modeled using the coefficient-of-performance (COP). The refrigeration power ($P_{refrigeration}$) can be expressed in terms of the COP and the COP of the Carnot cycle with Equation 2.5 and 2.6.

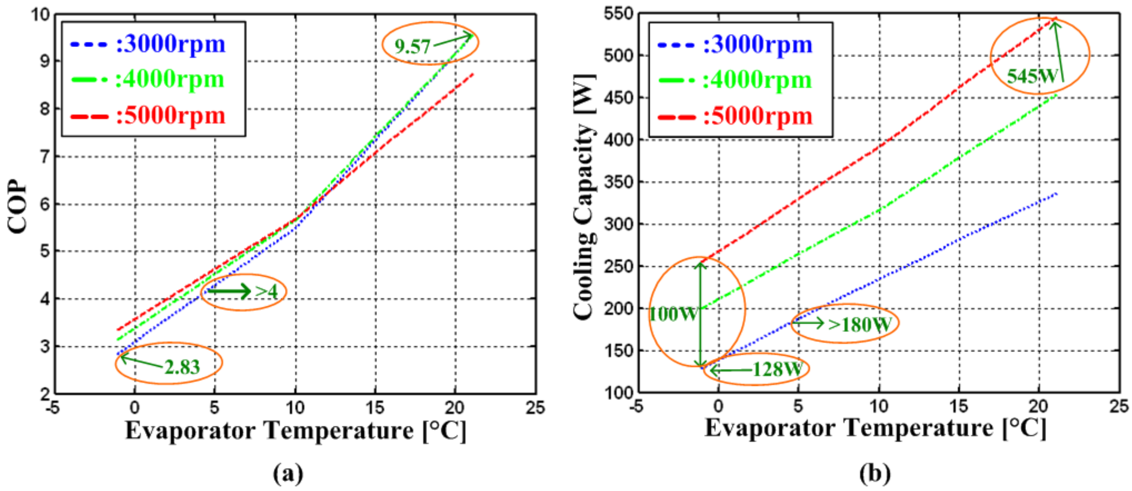


Figure 5-2: (a) COP and (b) Cooling capacity versus evaporator temperature at various compressor speeds for condenser temperature of 27°C.

Detailed performance of the compressor may be mapped as shown in Figure 5.2. These curves were adapted from [55] and indicate COP and cooling capacity as a function of evaporator temperature and compressor speed for a given condenser temperature. The system is optimally designed to function in the evaporator temperature ranges of -1 to 21°C. At these temperature ranges, the cooling capacity and COP of the compressor varies from 128 to 545W and 2.83 to 9.57, respectively at the condenser outlet of 27°C while operating at the compressor speed range of 3000 to 5000 RPM. As may be expected, COP increases with T_{evap} but stays relatively constant across different compressor speed. Also, the cooling capability increases with compressor speed. For example, COP is approximately 3 at T_{evap} of 0°C, and >100W of extra cooling capacity can be obtained by increasing the speed to 5000 from 3000 RPM at the expense of compressor power. COP in excess of 4.0 appears to be obtainable for >4°C for the maximum lifts up to 180 W at 3000 RPM. This performance far exceeds the power requirement of single-chip packages in high-performance systems and hence can potentially enable multi-processor servers.

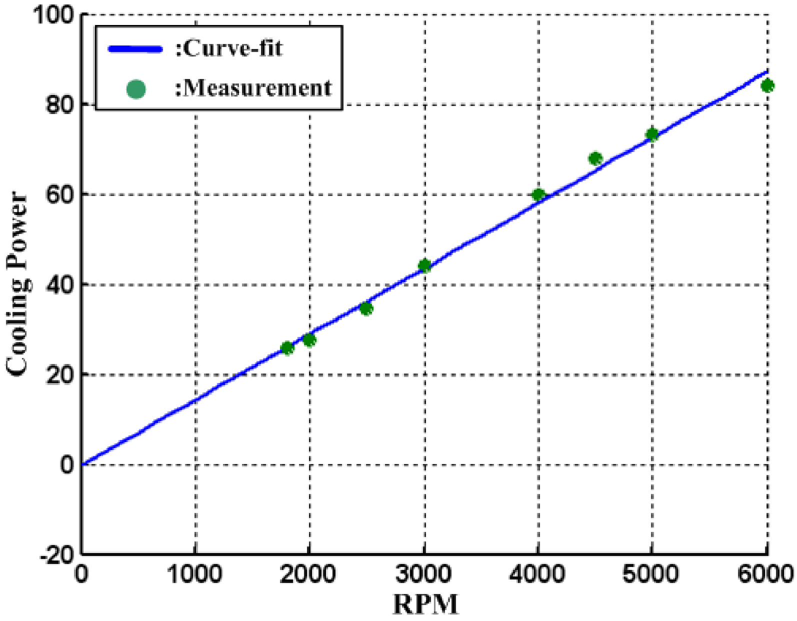


Figure 5-3: Cooling power versus RPM.

The choice of compressor speed is critical for achieving optimal power performance and is dependent on the system's range of electric power during operation. Figure 5.3 shows measured data and extrapolated curve of the compressor power at different speeds. Our hundred-watt range compressor is limited to the minimum speed of 1800 RPM while consuming 26W, implying that for a COP of 3 at T_{evap} of 0°C, the refrigeration system has cooling capacity of 78W. This lower limit indicates that our refrigeration system is overqualified for electric power of <78W and would result in excessive but unavoidable cooling cost. Assuming the compressor is linearly scalable (extrapolated curve in Figure

5.3), we explore the optimal operating temperatures and the amount of total power reduction at reduced temperatures at different operating conditions where the detail of the analysis is presented in Section 5.3.

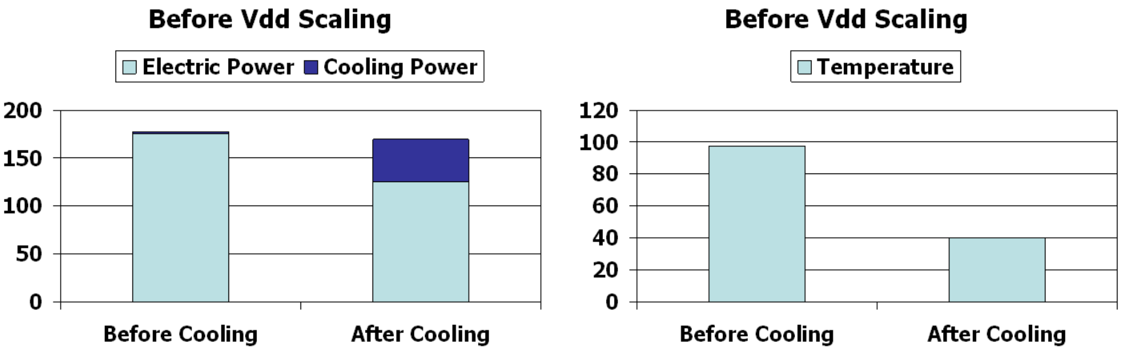
5.3 Experimental Results

Parameter	
$T_{junction0}$ [°C]	97
V_{dd0} [V]	1.25
Operating Frequency [GHz]	3.64
Leakage Percentage [%]	30
$P_{electric0}$ [W]	175.4
$T_{ambient}$ [°C]	22
P_{fan} [W]	1

Table 5.1: Performance summary of the microprocessor using a conventional forced-air convection cooling.

In this section, we show the impact of cooling a microprocessor using a highly-efficient miniature-scale refrigeration system using experimental results. Performance summary of the microprocessor using a conventional forced-air convection cooling is summarized in Table 5.1. The microprocessor operates at 3.64 GHz and dissipates 175.4W of maximum power ($P_{electric0}$) using a 1.25V power supply at 97°C ($T_{junction0}$). The $P_{leakage}/P_{electric}$ is 30% at this operating point, and the ambient temperature is 22°C. The microprocessor is directly cooled by the heatsink with fan forcing air flow over the heatsink. The fan forcing air flow over the heat sink dissipates approximately 1W of power.

Similar to our previous analysis in Chapter 3, the total power is analyzed at different temperature ranges while keeping the performance constant. Two different scenarios are tested with the refrigeration system. The first test maintains the supply voltage to 1.25V and the second test scales down the supply voltage to trade the improved performance for a reduction in power.



(a) (b)

Figure 5-4: Measurement results of the refrigerated cooling system before V_{dd} scaling.

The result of the first test is shown in Figure 5.4. Refrigerated cooling allows $T_{junction}$ to reach <40°C. Due to the exponential decrease in the leakage power at lower temperatures, the total $P_{electric}$ is reduced to 125W, reflecting a $P_{leakage}/P_{electric}$ ratio of 30%. Experimental results show that T_{evap} is around 0 ~ 5°C. The compressor operates at the speed of 3000 RPM to reject 125W of power. Cooling cost of the compressor running at this speed is approximately 44W leading to a total power of 169W which is only a modest amount of power savings.

A second test shows how scaling down the supply voltage can extract additional reduction in power as seen by Figure 5.5. By reducing the V_{dd} to 1.12V, an extra 28W of power reduction is obtained, which lead to total $P_{electric}$ of 97W. At this operating point, the compressor runs at 2500 RPM while consuming 35W of compressor power. The total power of 133W is obtained. This performance is 25% better than the non-cooled reference point. Note that power savings from supply scaling is compounded in cooling systems. Total amount of power savings can increase further if the compressor efficiency is improved.

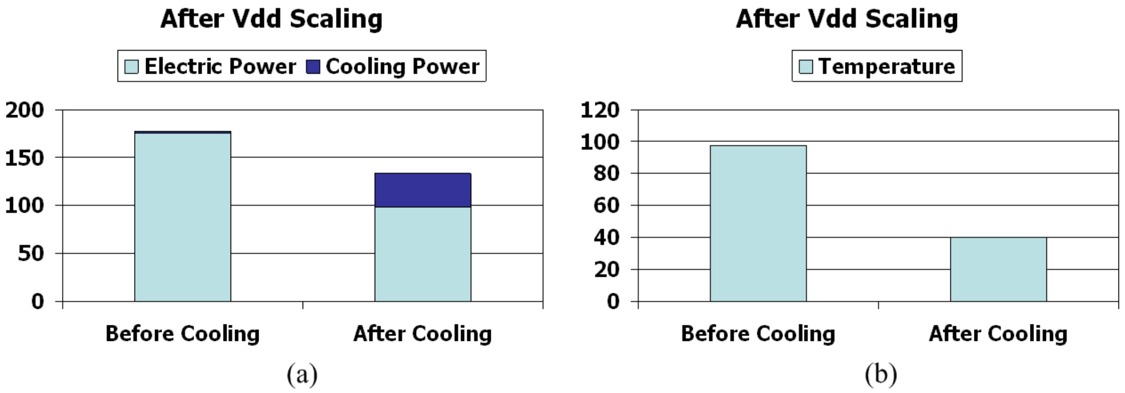


Figure 5-5: Measurement results of the refrigerated cooling system after V_{dd} scaling.

The most significant effect from cooling is the significant decrease in the leakage power shown in Figure 5.5. Note that the power savings is roughly proportional to the percent of leakage, which agrees with the analysis derived in Chapter 3. This measurement result implies that higher amount of power improvement can be obtained when the system power is dominated by leakage power. The power savings from voltage scaling is essentially offset by the cooling power.

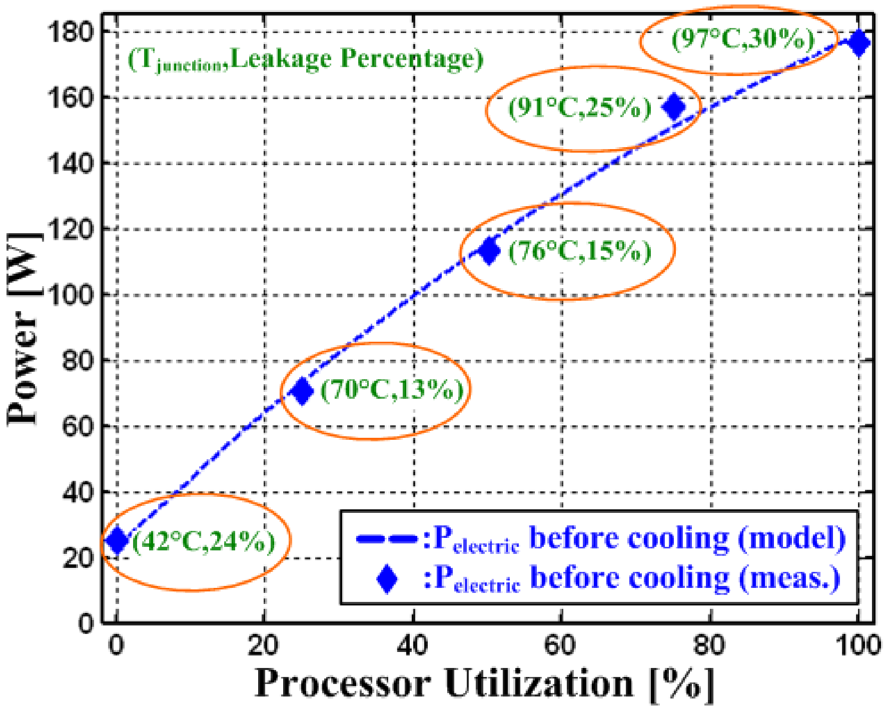


Figure 5-6: Power and operating junction temperature ($T_{junction}$) at different utilization percentage before refrigeration. (x,y) correspond to operating $T_{junction}$ [°C] and the percent of leakage power [%], respectively.

With different applications, the processor utilization percentage or activity factor changes. The electric power and the percentage of electric power due to its leakage and active components depend on the processor utilization and the operating junction temperature of the processor where the junction temperature of the processor is determined by the total power density and the ambient temperature through a given thermal resistance. For our particular microprocessor, $P_{electric}$ and $T_{junction}$ vary from 25W to 175.4W and from 42°C to 97°C, respectively, as shown in Figure 5.6. The figure also shows changes in leakage power at different processor utilization. The processor operating under these electronic profiles is cooled using our refrigeration system.

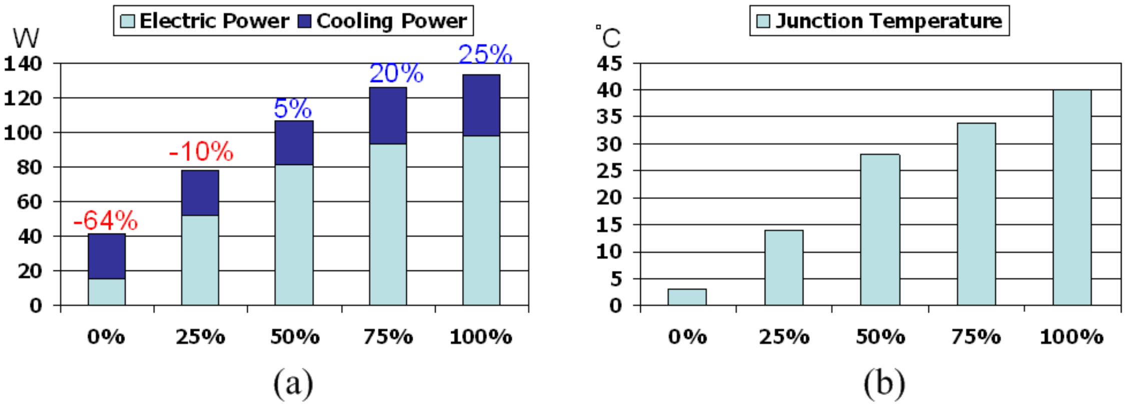


Figure 5-7: Measurement results of refrigeration cooling at different processor utilization.

The results of refrigeration cooling across different utilization percentage are shown in Figure 5.7. Note that the amount of electric power savings from cooling increases with the percentage of leakage power. Total power saving of 25% and 20% is obtained when the activity factor is at 100% and 75%, respectively. As described in Section 5.2, the efficiency of the compressor degrades significantly as the amount of required cooling capacity decreases. Hence, removing heat at below 50% utilization ratio results in total power consumption that is dominated by excessive cooling power. Our measurement results show that overall power penalty of 10 and 64% is obtained when the activity factor is at 25% and 0%, respectively.

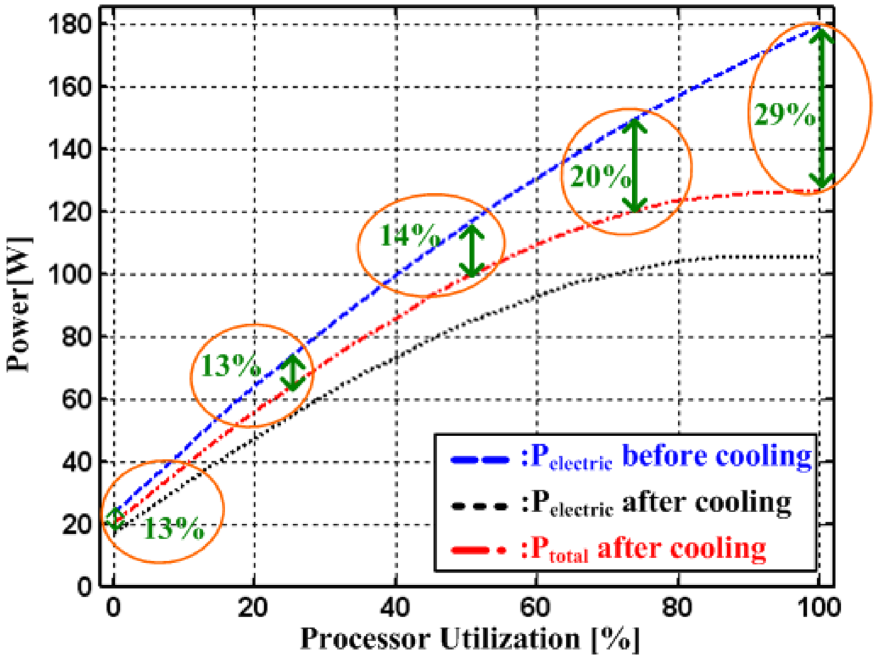


Figure 5-8: Result of the simulation of the total power at different processor utilization using the model from Section 2.1.

Since it is possible for a compressor to operate across a wider range of cooling capacity, we investigate the amount of total power saving at that reduced temperatures for different processor utilization assuming our refrigeration system has high COP across a broad range of cooling capacity. The result of this analysis is shown in Figure 5.8. The analysis shows that >13% of total power is saved across the entire range of processor utilization.

5.4 Summary

What has been demonstrated here is a refrigeration system that can not only lead to substantial power-performance improvement for the electronics, but the overall system performance improves. The total power savings is 25% when the processor is running at 100%. The amount of power saving depends on the amount of total electric power and the ratio of leakage to total electric power. The results of the experiments in this chapter indicate that enhancements of compressors that maintain high COP across the wide range cooling capacity is required in order for the refrigeration system to be used in electronics cooling for power minimization.

It is important to mention that while our analysis 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 simultaneously cool entire racks of servers by routing the coolant through the rack. We illustrate this potential configuration in Chapter 6.

CHAPTER 6

Effects of Cooling on Workload Management in HPCs

In this, we present an energy-efficient workload scheduling methodology for high-performance computing (HPC) servers under an actively-cooled environment where the purpose of the HPC servers is towards computation-intensive applications. Using highly efficient miniature-scale refrigeration systems for electronic cooling, we illustrate the potential of power optimization of multi-core multi-processor systems based on different workload scheduling methodologies 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 ≥29% of power reduction over the non-refrigerated design across the entire range of utilization levels. 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.

The chapter is organized as follows. In Section 6.1, we characterize the performance of a multi-core processor at reduced temperatures. Based on the characteristics, measurements, and the system-level model from Section 2.1, Section 6.2 and 6.3 describe an energy-aware workload scheduling methodology. Section 6.4 concludes with the main contributions of this chapter.

6.1 Multi-Core Processor Under the Actively Cooled Environment

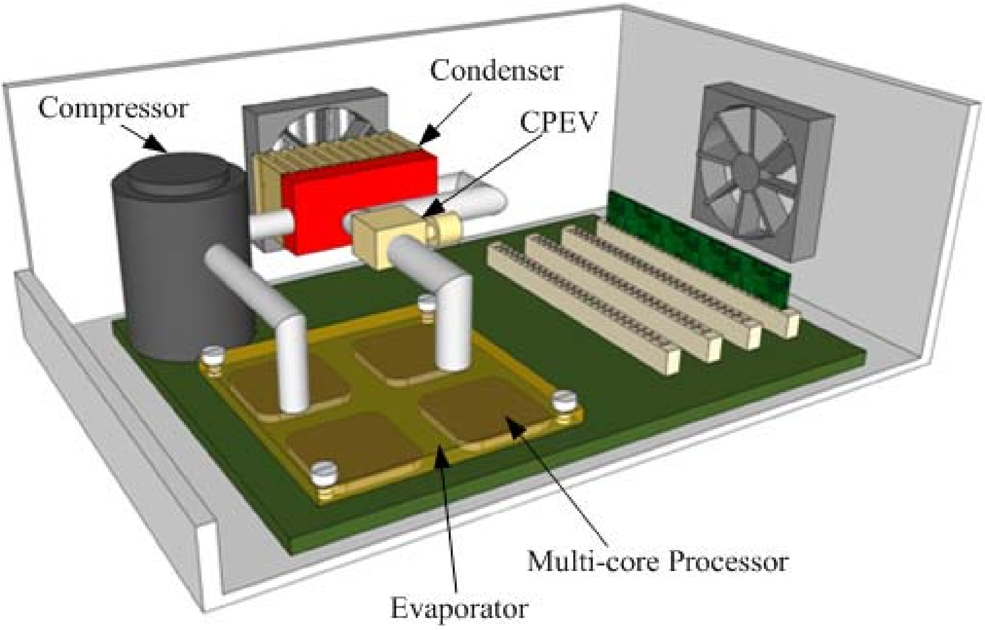
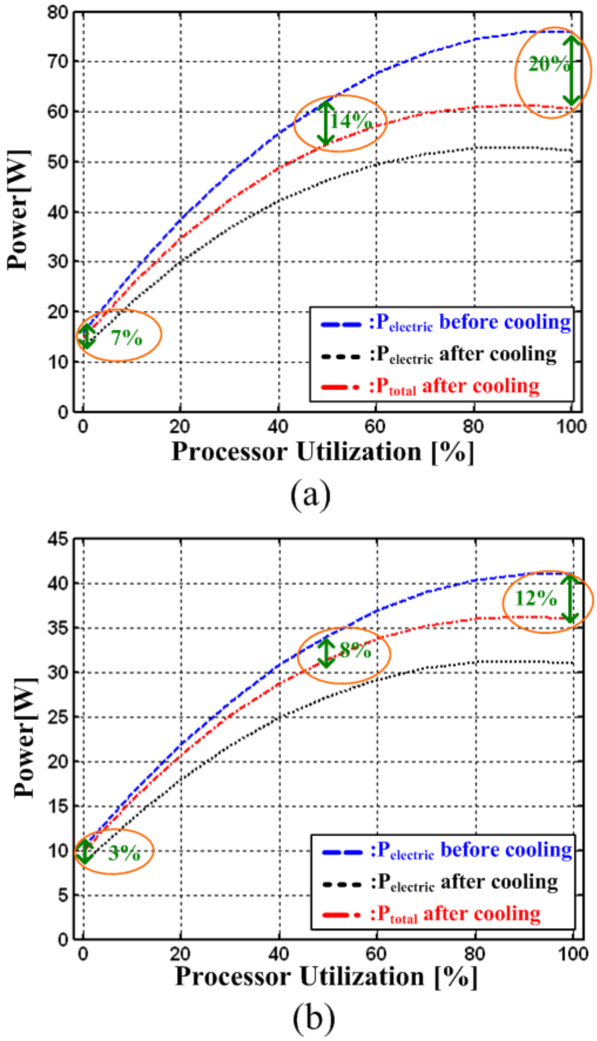


Figure 6-1: Configuration of a multi-processor computing server unit with a refrigerated-cooling.

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 Chapter 5. 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 Figure

6.1. One can apply a larger cooler used for a rack of servers with potentially better power efficiency. The results discussed in this dissertation can be directly applied. We characterize the power performance of a 4-core processor at different operating conditions using this refrigeration system. The results are used to build an empirical model of the 4-core processor operating at reduced temperatures and applied to multi-core multi-processors.

Furthermore, 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 Figure 5.8 and Figure 6.2. As can be seen, the total power savings of at least 3, 7 and 12 percent can be obtained across the entire range of processor utilization for 1, 2, and 4 cores respectively. The 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.



(a)

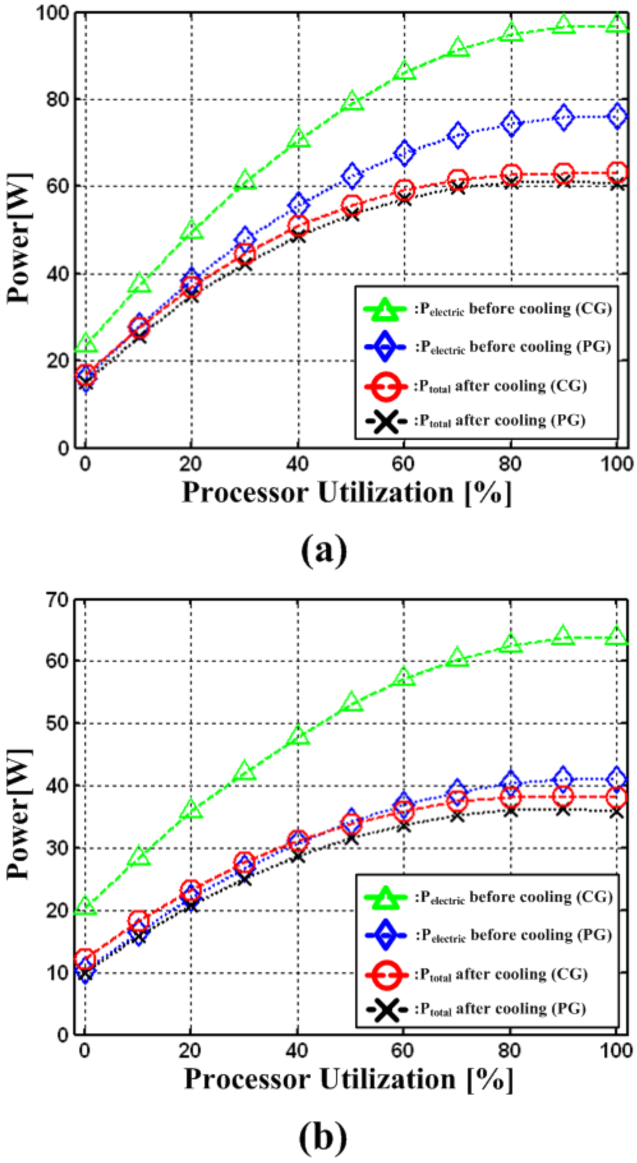
(b)

Figure 6-2: Measured and corresponding empirical model of power consumption before and after cooling across different processor utilization levels when (a) 2, or (b) 1 core out of the 4-core processor is powered up. The associated power savings after electronic cooling is also shown in the figure.

On the other hand, a simpler way to stop a core with minimal response time degradation is to clock gate (CG) the core [56]. This power saving 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 eliminates this problem since leakage power depends exponentially on temperature.

The result of incorporating reduced temperature is shown in Figure 6.3. 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.

By using the results from PG and CG, the next section discusses how workload scheduling can optimize the power dissipation of a multi-core, multi-processor system under different dynamic load variations.



(a)

(b)

Figure 6-3: 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.

6.2 Workload Scheduling Methodology

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 systems with 4-cores per processor under the actively cooled environment. In particular, we are interested whether active-cooling change the conventional way of assigning workloads.

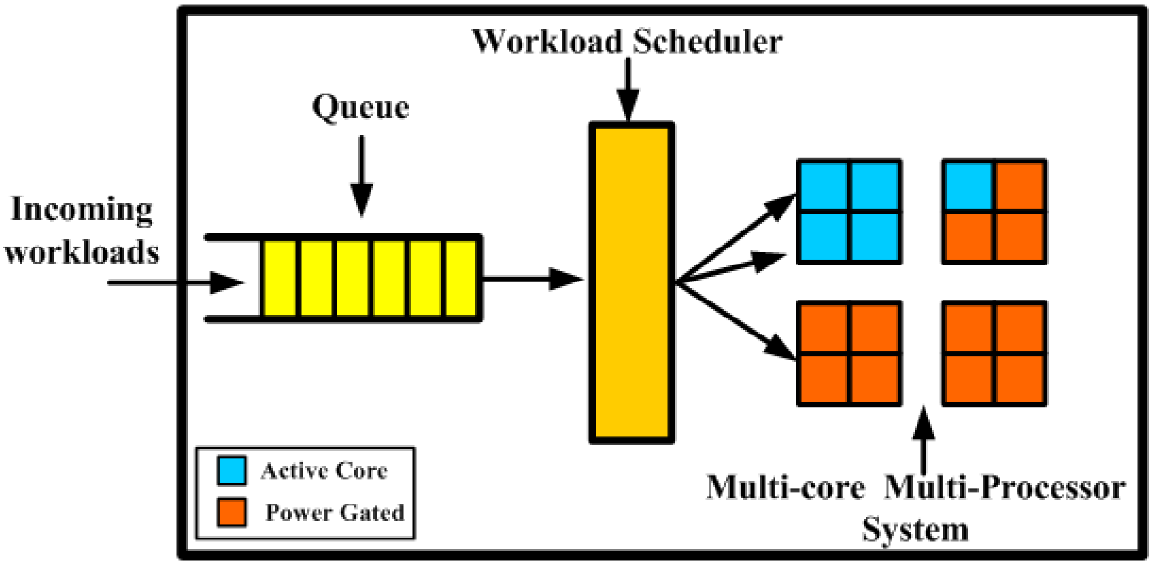


Figure 6-4: Generic workload scheduling management for multi-core multi-processor computing system.

Figure 6.4 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 400% utilization level where each core is responsible for 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.

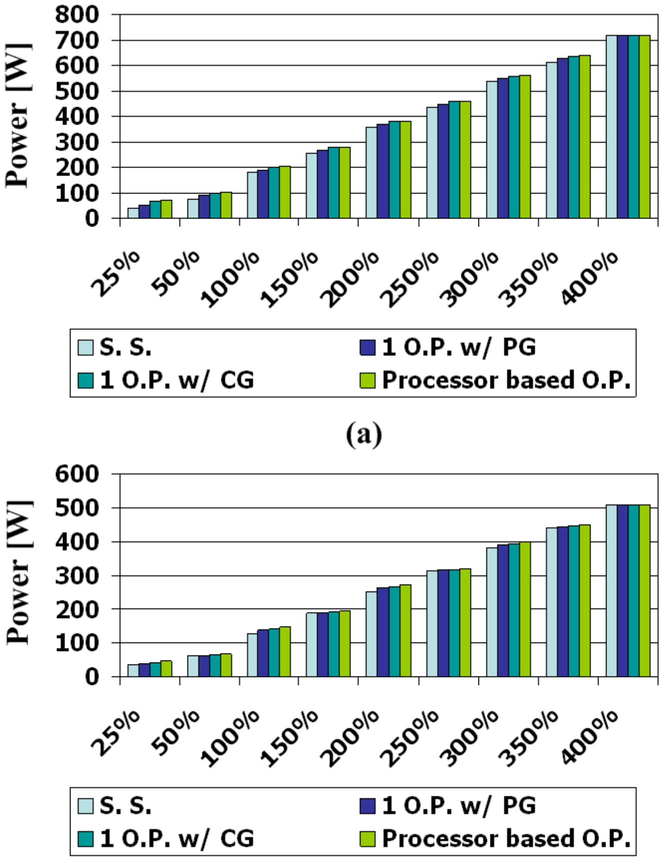
***1 core over-provision with CG (1 O.P. w/ CG)*:** Similar to *1 Core over-provision with PG* but uses CG for the core stopping mechanism.

***Processor based over-provision*:** 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 Figure 6.5. 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 Figure 6.6.

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 100% to 125%, 200% to 225%, and 300% to 325%.



(a)

(b)

Figure 6-5: Power at different utilization (a) before and (b) after electronic cooling for different methodologies.

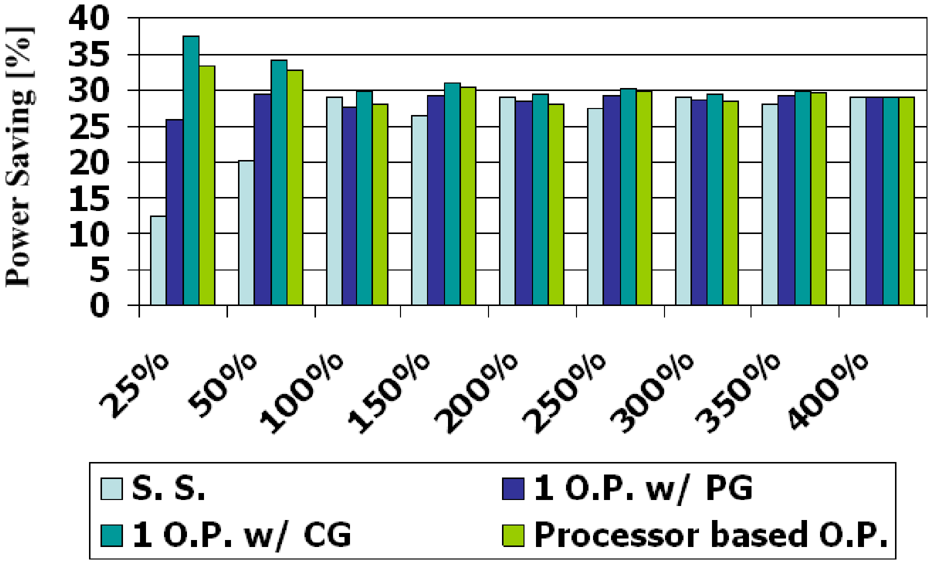


Figure 6-6: Associated power savings at different utilization level from electronic cooling for different workload assignment methodologies.

Next, we propose a new way of assigning workloads under refrigerated cooling and the approach is described in Figure 6.7. 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*. An example of the approach is shown in Figure 6.7; given a workload that requires 4 cores at 100% utilization, the workload scheduling is such that 4 cores are assigned equally to 2 processors. Figure 6.8 summarizes the results. 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 127W and

199W 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.

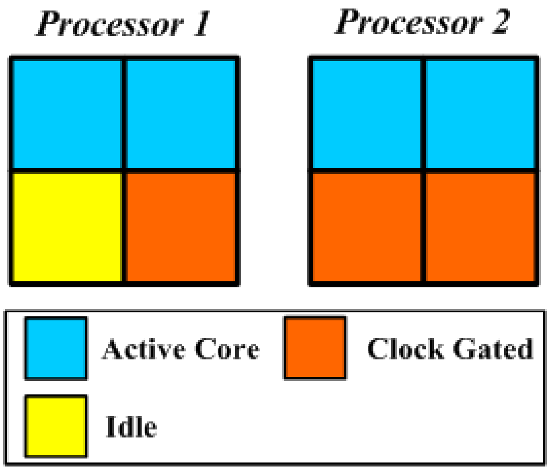


Figure 6-7: Proposed methodology.

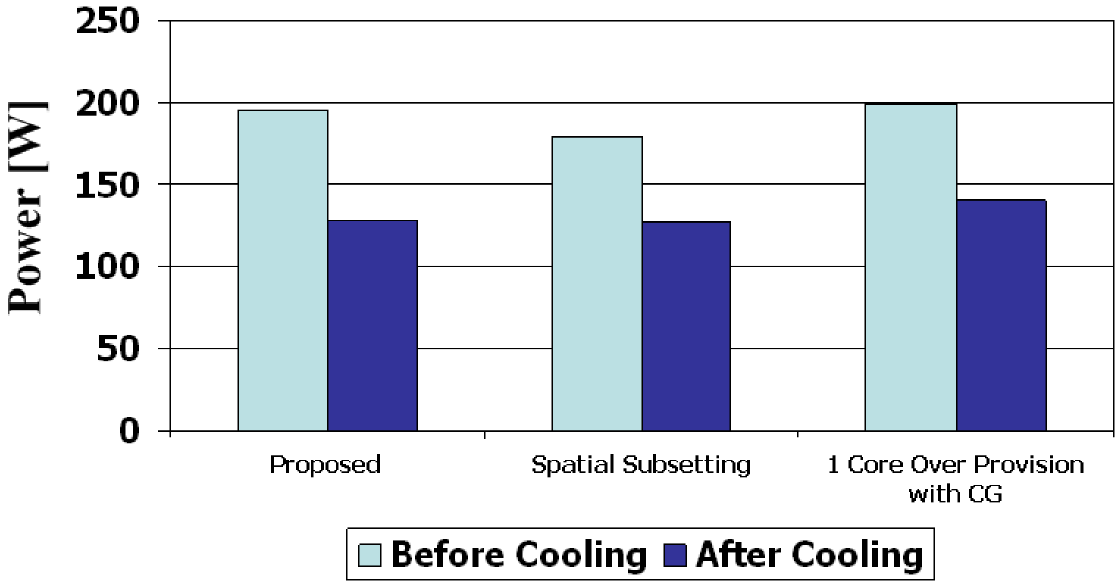


Figure 6-8: Proposed methodology compared with (B) *spatial subsetting* and (C) *1 core over-provision with CG*.

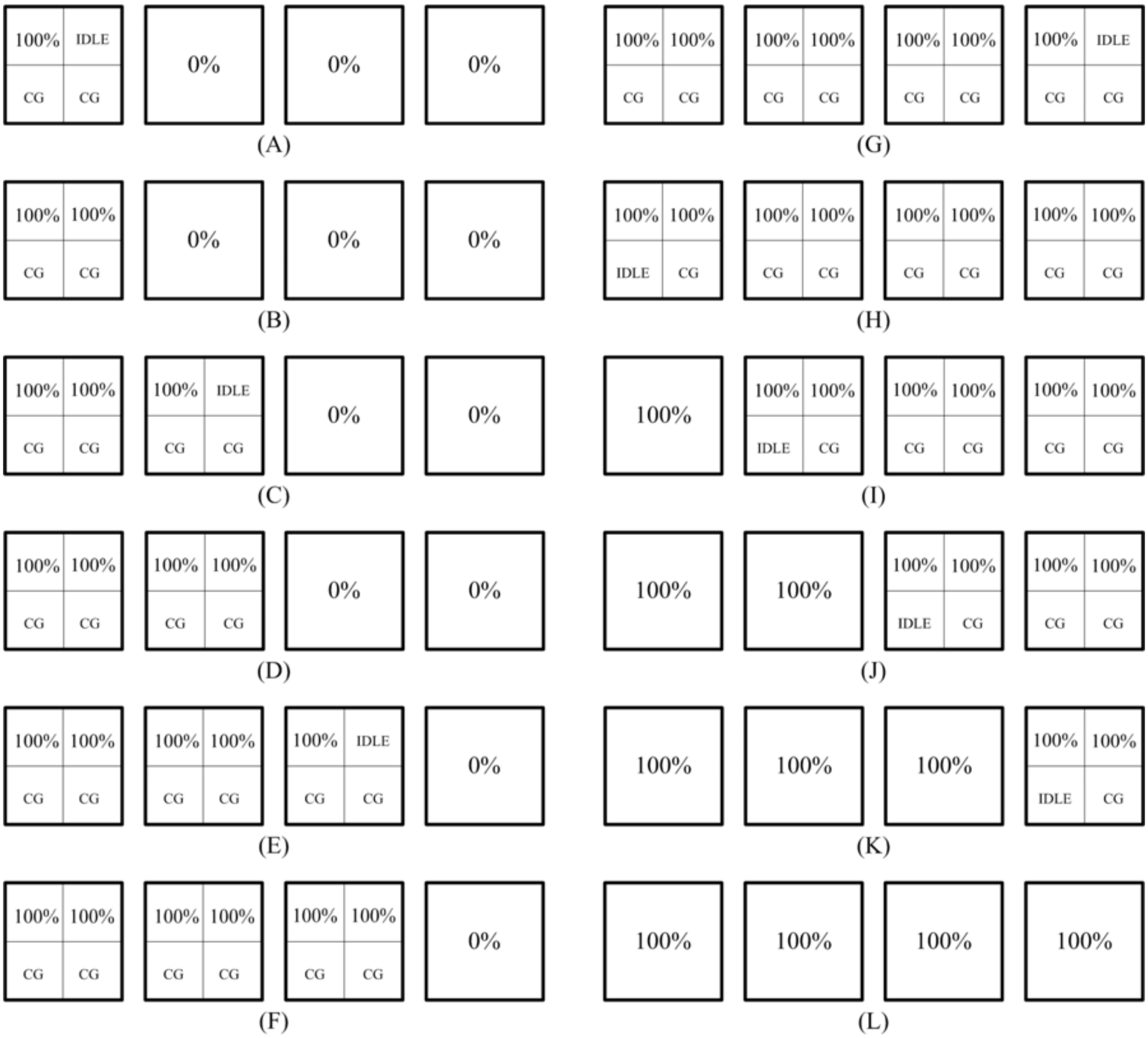
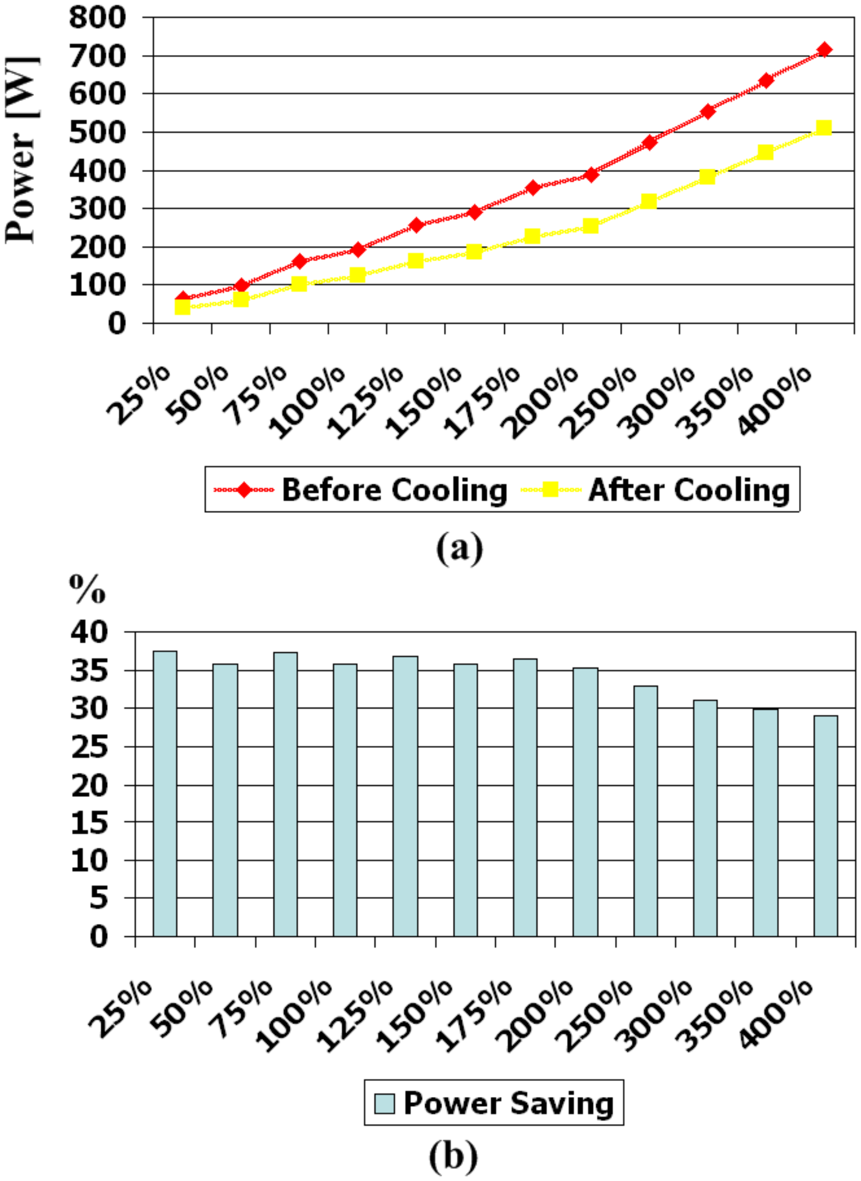

100% IDLE CG CG | 0% | 0% | 0%
(A)
100% 100% CG CG | 100% 100% CG CG | 100% 100% CG CG | 100% IDLE CG CG
(G)
100% 100% CG CG | 0% | 0% | 0%
(B)
100% 100% IDLE CG | 100% 100% CG CG | 100% 100% CG CG | 100% 100% CG CG
(H)
100% 100% CG CG | 100% IDLE CG CG | 0% | 0%
(C)
100% | 100% 100% IDLE CG | 100% 100% CG CG | 100% 100% CG CG
(I)
100% 100% CG CG | 100% 100% CG CG | 0% | 0%
(D)
100% | 100% | 100% 100% IDLE CG | 100% 100% CG CG
(J)
100% 100% CG CG | 100% 100% CG CG | 100% IDLE CG CG | 0%
(E)
100% | 100% | 100% | 100% 100% IDLE CG
(K)
100% 100% CG CG | 100% 100% CG CG | 100% 100% CG CG | 0%
(F)
100% | 100% | 100% | 100%
(L)

Figure 6-9: Proposed methodology across different utilization levels.

To be complete, we show the proposed workload scheduling methodology for different utilization levels in Figure 6.9 Since power-up events are necessary when a new processor is brought online, three power-up transition delays are necessary. 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. Figure 6.10 plots the power dissipation and the percentage savings before and after cooling for each of the conditions shown in Figure 6.9.



(a)

(b)

Figure 6-10: (a) Power consumption at corresponding utilization levels of Figure 6.9. (b) The associated power saving from electronic cooling.

6.3 Assignment of Workload Based on Target 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. 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 Figure 6.11. Figure 6.12 shows the execution velocity for each of these 3 workload scenarios. 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.

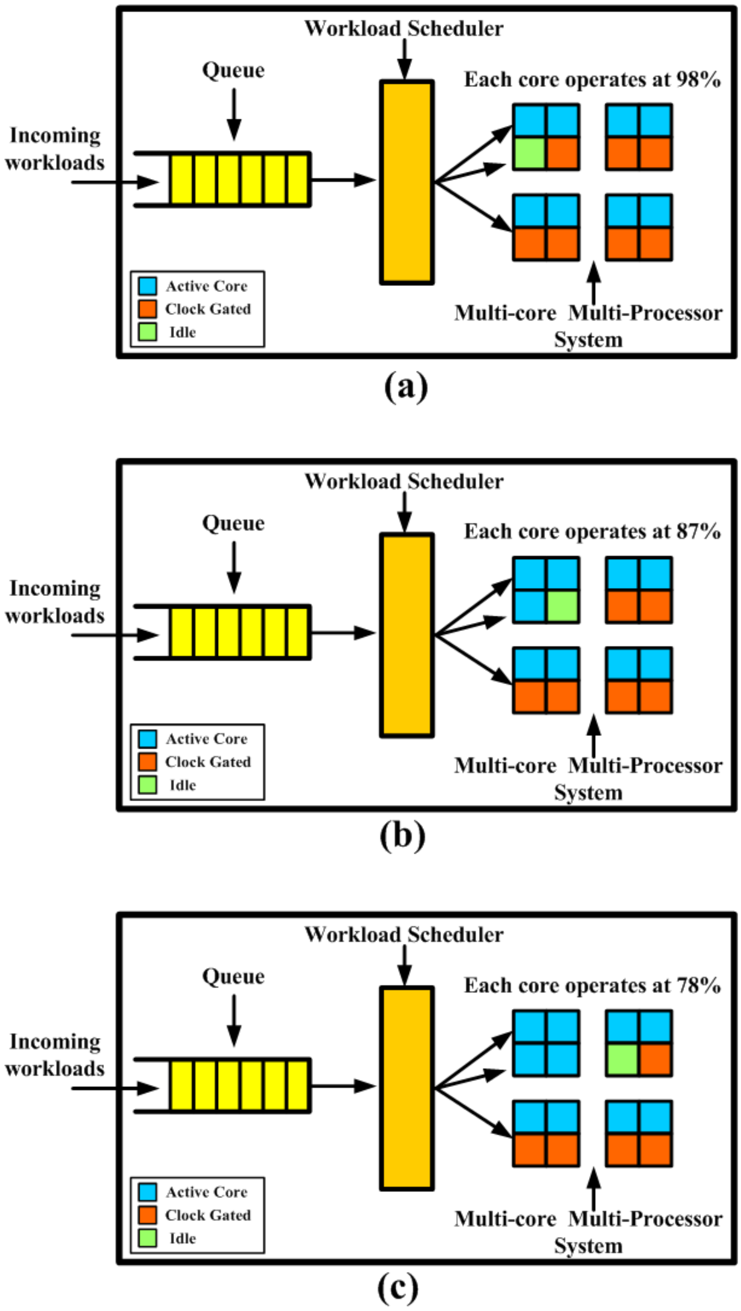


Figure 6-11: Required utilization level across different number of processors (M) for the proposed methodology.

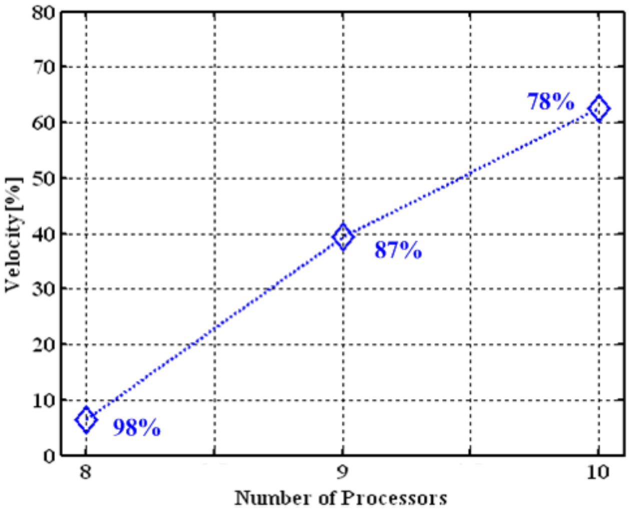


Figure 6-12: Execution velocity vs. number of processors. Number in the figure represents required utilization level.

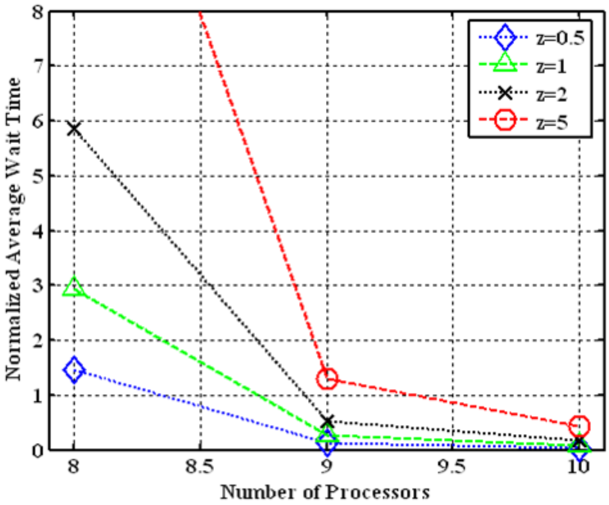
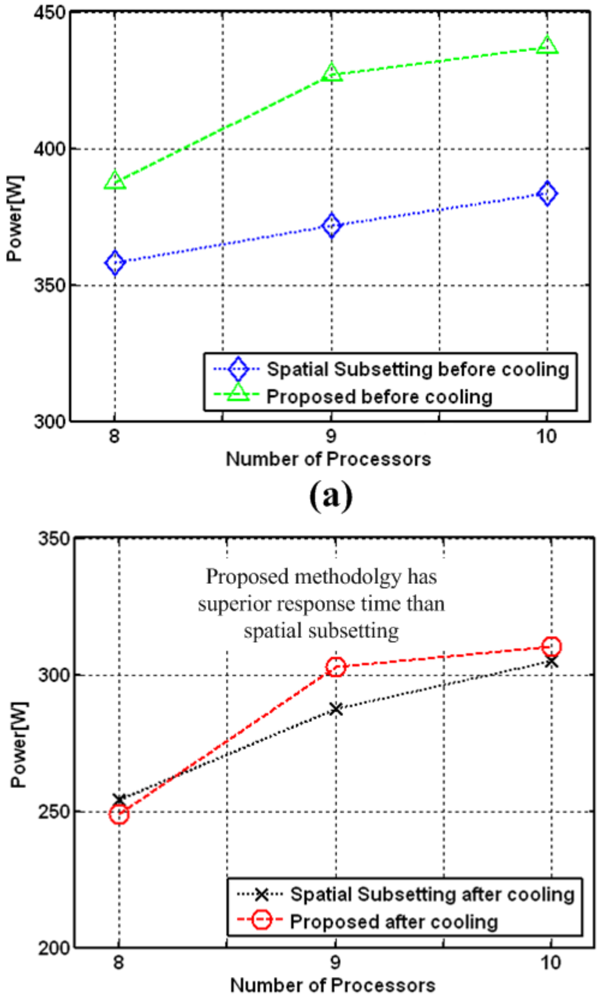


Figure 6-13: Normalized average wait time vs. number of processors as function of workload variability z=0.5, 1.0, 2.0, 5.0.

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. We consider $0 \leq$ E[W] ≤ 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 Figure 6.13).



(a)

(b)

Figure 6-14: Power consumption vs. number of processors (a) before and (b) after cooling.

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 Figure 6.14. 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.

6.4 Summary

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, but still maintains 29% of power reduction.

CHAPTER 7

Conclusion

This dissertation analyzes the tradeoffs of employing active electric cooling systems to address the power and thermal problems of high-performance computing systems and to demonstrate an approach that leads to overall system power improvement that includes the cost of cooling power. Our approach of using such cooling systems is attractive with deeper technology scaling due to the increase in power density and leakage power. Previous studies have considered the design of cooling system as a separate design space without taking account of power/thermal characteristics of the electronics. In this dissertation, we emphasize the importance of incorporating both the electronic and refrigeration systems in order to realize overall system power improvement as there exists a strong interdependence between the two systems. The first section in this chapter summarizes the techniques proposed in this dissertation, and Section 7.2 describes some interesting challenges of future cooling-assisted computing systems design.

7.1 Contributions

The proposed system-level modeling expedite the analysis of optimal temperature operating points, the amount of total power reduction at reduced temperatures, and the dependence of the refrigerated performance on the power profile of the electronics. Our analytical result of using two-phase cooling methods shows that the amount of power savings is roughly proportional to the ratio of leakage power to total power due to the sensitivity of leakage power to temperature. Our analysis also shows that the optimal target temperature lies within the range of domestic refrigeration systems, which gives the possibility of using widely-available and well-established domestic refrigeration technology for electronic cooling. Adapting the operating temperature of refrigeration system to the operation of the electronics can lead to better power performance, but the difference is small as compared to using a constant temperature due to the poor refrigeration efficiency as the temperature deviates from the target refrigerator design. Unless the refrigeration system has a broad range with high COP near the targeted temperature, the benefits of adapting the optimal temperature may not be worth the extra design overhead. Finally, based on various parametric analyses, we show that the overall system performance is most sensitive to the performance of refrigeration systems and enhancements to refrigeration systems at domestic freezer temperatures can further lead to greater power savings.

Next, the feasibility of using a different cooling mechanism, eTEC, is explored to solve power and thermal problems. Analysis indicates that the performance primarily relies on the amount of leakage power and the ratio of leakage to total electronic power. While providing advantage of fully integrated solution, power savings using eTECs are not as significant as cooling using other means such as a vapor compression cycle at the chip level. Nevertheless, cooling the right functional block can still yield a modest power improvement. Enhancing the performance of the eTEC provides a much more compelling performance improvement in terms of power (>10%). Even without substantial power improvements, by cooling specific functional units that are potential hot-spots, IC reliability improves substantially. Finally, since the amount of leakage power increases with technology scaling, using eTECs may be a possible power/thermal solution for future electronics.

A miniature-scale refrigeration system for electronic cooling is developed to experimentally demonstrate the amount of total power saving for processors at various operating conditions. Our experiment shows that the processor dissipating 175.4W of power with 30% electronic leakage power and operating at 97°C offers a 25% total system power reduction over the non-refrigerated design with junction temperature operating at <40°C. Not only the measurement results validate our system-level model, but this experiment is the *first demonstration* of active cooling that lead to reduced total wall power. The results from the experiments also indicate that enhancements of compressors that maintain high COP across the wide range cooling capacity is required in order for the refrigeration system to be used in electronics cooling for power minimization.

In addition, we illustrate the potential of simultaneously cooling multi-core multi-processor systems. This potential configuration led to the study of energy-efficient workload scheduling methodology for HPC servers using our miniature scale refrigeration system. While leveraging the benefits of clock gating at reduced temperatures, the proposed methodology results in total power consumption that is comparable to the *spatial subsetting* scheme but with superior response time. Our actively cooled system results in $\geq 29\%$ power reduction over the non-refrigerated design across the entire range of utilization levels of multi-core multi-processor system. Furthermore, combining our proposed methodology with the approximation formulas for a G/G/m-model, we show that we can reach an optimal agreement between high utilization of the processors (energy-conscious provisioning) and the target SLA requirements.

7.2 Future Work

An interesting research topic is to explore different feedback-driven control solutions that provide capability to adapt to diverse environment, workload, and user constraints. 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.

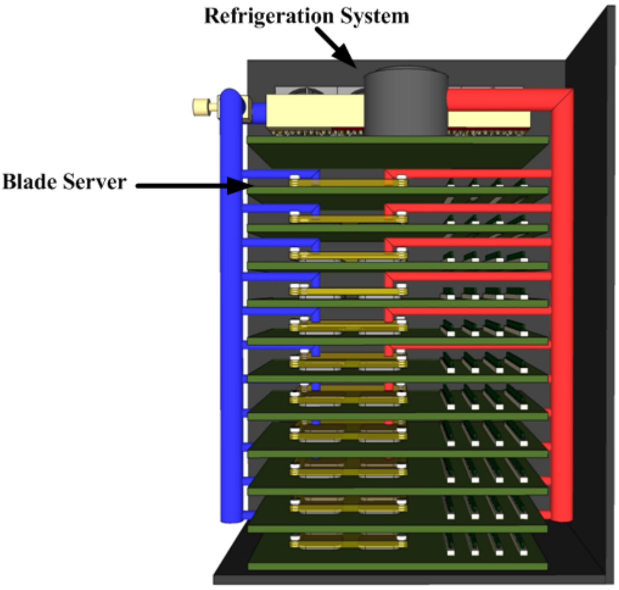


Figure 7-1: Configuration of a server rack with refrigerated-cooling.

Note that 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 by routing the coolant through the rack as shown in Figure 7.1. This dissertation has proposed a feasible workload scheduling policy for multi-core multi-processor server system at reduced temperatures that offers lower power consumption compared to non-cooled design. While the results discussed in this dissertation 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, as our proposed techniques in this dissertation can directly be applied to datacenters as shown in Figure 7.2, thorough understanding of the strong coupling between refrigerated server racks and CRAC units is needed for future research.

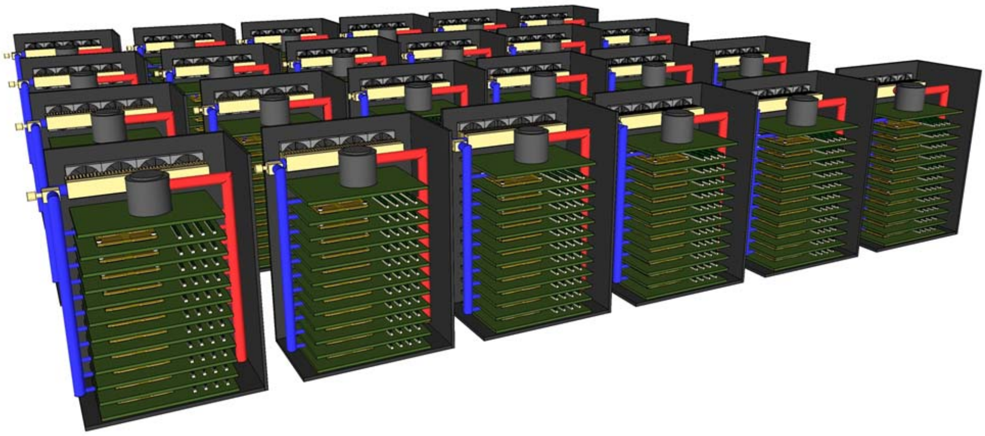


Figure 7-2: Future datacenters with refrigerated server racks.

Bibliography

[1] U.S. EPA. Report to congress on server and data center energy efficiency. In U.S. Environmental Protection Agency, Tech Report, 2007.

[2] 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.

[3] J. Humphreys and J. Scaramella, " The impact of power and cooling on data center infrastructure," Market Research Report, IDC, 2006.

[4] Tschudi, et al., "Data Centers and Energy Use – Let's Look at the Data", *ACEEE*, 2003.

[5] Lawrence Berkeley National Labs, Benchmarking: Data Centers, Dec 2007.

[6] M. Patterson, "The effect of data center temperature on energy efficiency," *Proceedings ITHERM*, pp. 1167-1174, 2008.

[7] 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.

[8] 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.

[9] 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.

[10] D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: A framework for architectural-level power analysis and optimizations," In I*nternational Symposium on Computer Architecture*, June 2000.

[11] 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.

[12] 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.

[13] R. McGowen, C.A. Poirier, C. Bostak, J. Ignowski, 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.

[14] S. Heo, K. Barr, and K. Asanovic, "Reducing power density through activity migration," *International Symposium on Low Power Electronics and Design*, August 2003.

[15] D. Copeland, "64-bit Server Cooling Requirements," *IEEE SEMI-THERM* Symposium. 2005.

[16] R. Mahajan, C. P. Chiu, and G. Ghrysler, "Cooling a Microprocessor Chip," in *Proceedings of the IEEE*, vol. 94, no. 8, August 2006.

[17] A.G. Agwu Nnanna, "Application of refrigeration system in electronics cooling," *Applied Thermal Engineering*, vol. 26, pp. 18-27, 2006.

[18] International Technology Roadmap for Semiconductors (ITRS), 2010 edition, http://public.itrs.net/

[19] R. C. Chu, R. E. Simons, M. J. Ellsworth, R. R. Schimidt, and V. Cozzolino, "Review of Cooling Technologies for Computer Products," *IEEE Transactions on Device and Material Reliability*, vol. 4, no. 4, pp. 568-585, December 2004.

[20] P. E. Phelan, V. A. Chiriac, and T. T. Lee, "Current and Future Miniature Cooling Technologies for High Power Microelectronics," *IEEE Transactions on Components and Packaging Technologies*, vol. 25, no. 3, pp. 356-365, September 2002.

[21] S. Trutassanawin, E. Groll, V. Garimella, and L. Cremaschi, "Experimental Investigation of a Miniature Scale Refrigeration System for Electronics Cooling," *IEEE Transactions on Components and Packaging Technologies*, vol. 29, no. 3, pp. 678-687, September 2006.

[22] L. Jiang, et al., "Closed-Loop Electroosmotic Microchannel Cooling System for VLSI Circuits," *IEEE Transactions on Components and Packaging Technologie*s, vol. 25, no. 3, September 2002.

[23] S. Krishnan, S. V. Garimella, G. M. Chrysler, and R. V. Mahajan, "Towards Moore's Law," *IEEE Transactions on Advanced Packaging*, vol. 30, no. 3, pp. 462-474, August,2007.

[24] G. J. Snyder, M. Soto, R. Alley, D. Koester, and B. Conner, “Hot Spot Cooling Using Thermoelectric Coolers,” *IEEE SEMI-THERM Symposium*, 2006.

[25] I. Chowdhury, et al., “On-chip cooling by superlattice-based thin-film thermoelectrics,” *Nature Nanotechnology*, 2009, vol. 4, pp. 235-238, April 2009.

[26] 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.

[27] S. Borkar, “Getting Gigascale chips: Challenges and opportunities in continuing Moore’s law,” *ACM Queue*, vol. 1, pp. 26-33, October 2003.

[28] G. M. Link, N. Vijaykrishnan, “Thermal Trends in Emerging Technologies,” Proceedings of the 7th International Symposium on Quality Electronic Design, pp. 625-632, March 27-29, 2006.

[29] V. Karkare, S. Gibson, and D. Markovic, “A 130uW, 64-Channel Spike-Sorting DSP Chip,” in IEEE Asian Solid-State Circuits Conference, pp. 289-292, November 2009.

[30] G. Gammie, A. Wang, M. Chau, S. Gururajarao, R. Pitts, F. Jumel, S. Engel, P. Royannez, R. Lagerquist, H. Mair, J. Vaccani, G. Baldwin, K. Heragu, R. Mandal, M. Clinton, D. Arden, and U. Ko, “A 45nm 3.5G Baseband-and-Multimedia Application Processor using Adaptive Body-Bias and Ultra-Low-Power Techniques,” *ISSCC Digest of Technical Papers,* pp. 258-259, February 2008.

[31] V. Ramadurai, H. Pilo, J. Andersen, G. Braceras, J. Gabric, D. Geise, S. Lamphier, and Y. Tan., “An 8Mb SRAM in 45nm SOI Featuring a Two-Stage Sensing Scheme and Dynamic Power Management,” *IEEE Journal of Solid-State Circuits,* vol. 44, no. 1. pp 155-162, January 2009.

[32] G. Gerosa, S. Curtis, M. D’Addeo, J. Bo, B. Kuttanna, F. Merchant, B. Patel, M. H. Taufique, and H. Samarchi., “A Sub-2 W Low Power IA Processor for Mobile Internet Device in 45nm High-k Metal Gate CMOS,” *IEEE Journal of Solid-State Circuits,* vol. 44, no. 1. pp 73-82, January 2009.

[33] G. K. Konstadinidis, M. Tremblay, S. Chaudhry, M. Rashid, P. F. Lai, Y. Otaguro, Y. Orginos, S. Parampalli, M. Steigerwald, S. Gundala, R. Pyapali, L. D. Rarick, I. Elkin, Y. Ge, and I. Parulkar., “Architecture and Physical Implementation of a Third Generation 65nm, 16 Core, 32 Thread Chip-Multithreading SPARC Processor,” IEEE Journal of Solid-State Circuits, vol. 44, no. 1. pp 7-17, January 2009.

[34] Sanyo Denki, Standard Fan SAN ACE 40 for 1U server, datasheet published at http://db.sanyodenki.co.jp/product_db_e/coolingfan/

[35] S. P. Gurrum, S. K. Suman, Y. K. Joshi, and A. G. Fedorov, “Thermal issues in next-generation integrated circuits,” *IEEE Transactions on. Device Materials Reliability*, vol. 4, no. 4, pp. 709-714, April 2004.

[36] Berchowitz D.M. "Miniature Stirling Coolers", *NEPCON - East '93*, Boston, MA, June, 1993.

[37] N. W. Lane, J. G. Wood, R. Z. Unger, “Free-piston Stirling Machine Commercialization Status atSupower,” International Stirling Engine Conference, 19-21 November 2003.

[38] Hitachi Appliances, Inc., http://www.hitachi-ap.com/index.html/

[39] Global Cooling, Inc., http://www.globalcooling.com/

[40] Li, Sheng, et al., “McPAT: An Integrated Power, Area, and Timing Modeling Framework for Multicore and Manycore Architectures,” in *IEEE MICRO*, pp. 469-480, 2009.

[41] A. Jain, et al., “A 1.2GHz Alpha Microprocessors with 44.8 GB/s Chip Pin Bandwidth,” *ISSCC Digest of Technical Papers*, pp.240-241, Feb. 2001.

[42] A. S. Leon, K. W. Tam, J. L. Shin, D. Weisner, and F. Schumacher, “A Power-Efficient High-Throughput 32-Thread SPARC Processor,” *IEEE Journal of Solid-State Circuits*, vol. 42, no. 1, January 2007.

[43] S. Rusu, S. Tam, H. Muljono, et al., “A Dual-Core Multi-Threaded Xeon Processor with 16MB L3 Cache,” *ISSCC Digest of Technical Papers*, pp. 118-119, February, 2006.

[44] 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,” *Proceedings of the 36th ACM/IEEE International Symposium on Microarchitecture*, December 2003.

[45] 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.

[46] 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.

[47] 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.

[48] 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.

[49] B Müller-Clostermann, "Using G/G/m-Models for Multi-Server and Mainframe Capacity Planning," *ICB Research Report*, no. 16, May 2007.

[50] D. M. Carlson, D. C. Sullivan, R. E. Bach, and D. R. Resnick, "The ETA-10 liquid-nitrogen-cooled supercomputer system," *IEEE Transactions on Electron Devices,* vol. 36, no. 8, pp. 1404-1413, Aug. 1989.

[51] I. Aller, K. Bernstein, U. Ghoshal, H. Schettler, S. Schuster, Y. Taur, and O. Torreiter., "CMOS Circuit Technology for Sub-Ambient Temperature Operation," *ISSCC Digest of Technical Papers*, pp. 214-215, Feb. 2000.

[52] M. J. Deen, "Digital Characteristics of CMOS Devices at Cryogenic Temperatures," IEEE Journal of Solid-State Circuits, vol. 24, no. 1, pp. 158-164, February 1989.

[53] S. L. Lin and K. Banerjee, "Cool Chips: Opportunities and Implications for Power and Thermal Management," *IEEE Transactions on Electron Devices*, vol. 55, no. 1, pp. 245-255, Jan. 2008.

[54] Rich Miller, http://www.datacenterknowledge.com/archives/2008/10/14/google-raise-your-data-center-temperature/

[55] Aspen Compressor, LLC, www.aspencompressor.com

[56] N. A. Kurd, J. S. Barkatullah, R. O. Dizon, T. D. Fletcher, and P. D. Madland, "A multigigaherz clocking scheme for Pentium 4 microprocessor," *IEEE Journal of Solid-State Circuits*, vol. 36, pp. 1647-1653, November 2001.

[57] Moran M, Shapiro H. Fundamentals of engineering thermodynamics, 3rd ed. New York: Wiley, 1996.

[58] J. Rabaey, A. Chandrakasan, and B. Nikolic, Digital Integrated Circuits: A Design Perspective; 2nd ed., 2003.