

Mitigating Dark-Silicon Problems Using Superlattice-based Thermoelectric Coolers

Francesco Paterna and Sherief Reda

School of Engineering, Brown University - 184 Hope St., Providence, RI 02912
emails: {francesco_paterna, sherief_reda}@brown.edu

Abstract—Dark silicon is an emerging problem in multi-core processors, where it is not possible to enable all cores simultaneously because of either insufficient parallelism in software applications or because of high-spatial power densities that generate hot-spot constraints. Superlattice-based thermoelectric cooling (TEC) is a promising technology that offers large heat pumping capability and the ability to target hot spots of each core independently. In this paper, we devise novel system-level methods that address the two main sources of dark silicon using superlattice TECs. Our methods leverage the TECs in conjunction with dynamic voltage and frequency scaling and number of threads to maximize the performance of multi-core processor under thermal and power constraints. Using an experimental setup based on a quad-core processor, we provide an evaluation of the trade-offs among performance, temperature and power consumption arising from the use of superlattice-based TECs. Our results demonstrate the potential of this emerging cooling technology in mitigating dark silicon problems and in improving the performance of multi-core processors.

I. INTRODUCTION

An emerging problem in modern processors is *dark silicon*, where it is not possible to enable all functional units of a processor simultaneously [6]. This limitation is arising from two causes: 1) high-density spatial power consumption and hot spots developed in a few units, and 2) the inefficient parallelization of some applications. Hot spots compromise the performance and reliability of semiconductor devices and as a result they must be prohibited [7].

While using dynamic voltage and frequency scaling (DVFS) enables the reduction of power consumption and consequently the magnitude of hot spots, this approach adversely impacts performance. Emerging superlattice-based thermoelectric coolers (TECs) promise to target individual hot spots directly at the micro-scale, alleviating the constraints on the performance of processors [5]. In contrast to traditional TECs, superlattice-based TECs offer high-density spatial heat pumping capabilities in a small form factor, making them ideal to integrate between the die of the processor and the heat spreader as illustrated in Figure 1.

In contrast to existing works which investigate physical issues or control issues at the micro-architectural level [1], [9], [4], we mainly focus on system-level methods that leverage superlattice TECs in conjunction with DVFS to mitigate the two main causes of dark silicon. The main contributions of this paper are as follows.

This work is partially supported by a NSF CAREER grant number 0952866.

978-3-9815370-0-0/DATE13/ © 2013 EDAA

- We propose system-level power and thermal models for multicore processors integrating superlattice TECs. These models estimate the impact of critical system-level settings, such as DVFS, power of the TECs, and number of application threads on performance, power consumption, and thermal characteristics.
- We formulate system-level management methods to identify the best settings to mitigate the main two causes of dark silicon. If the threads are heterogeneous with sufficient parallelism, then our method activates a large number of cores to increase throughput, while constraining the hot spot temperatures, and if the threads do not have enough parallelism, then our method boosts the DVFS settings of a fewer cores to increase throughput.
- Using measurements from a quad-core processor, we parametrize our models and validate their power and thermal results. The models are then used for experimentation. We demonstrate the effectiveness of our management methods using heterogeneous workloads from the SPEC CPU06 and PARSEC benchmarks.

The remainder of this paper is organized as follows. We briefly provide the background of superlattice TECs in Section II. In Section III, we propose our system-level thermal and power modeling techniques. In Section IV, we formulate optimization methods to mitigate dark-silicon problems in multi-core processors. We demonstrate the effectiveness of our methods experimentally in Section V. Finally, Section VI provides the main conclusions of this work.

II. BACKGROUND ON SUPER-LATTICE TECs

A thermoelectrical cooler (TEC) is a device based on the Peltier effect such that if a current I passes through a thermocouple, a quantity of heat is absorbed at one junction and

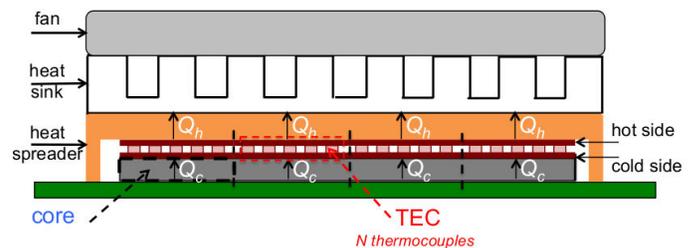


Fig. 1. Integration of per-core superlattice TECs with a multi-core processor.

rejected at the other, effectively pumping heat. The pumped heat depends on the current through the Peltier coefficient between the two junctions. The inverse of the Peltier effect is the Seebeck effect which states that if there is a thermal gradient through a thermocouple a voltage difference is developed proportionally through the Seebeck coefficient S [8].

Let's consider a TEC composed by N parallel p-n thermocouples, each thermocouple characterized by a length l , area A_p , and a thermal resistance $R_t = l/(2kA_p)$ and a electric resistance $R_e = 2\rho l/A_p$, where k and ρ are the thermal conductivity and electrical resistive respectively. Let's denote the temperature at the hot junction by T_h and that at the cold junction T_c , the heat absorbed at the cold side is given by Equation (1), while the heat rejected at the hot side is given by Equation (2).

$$Q_c(I, T_h, T_c) = N \left(S I T_c - \frac{T_h - T_c}{R_t} - \frac{1}{2} I^2 R_e \right) \quad (1)$$

$$Q_h(I, T_h, T_c) = N \left(S I T_h - \frac{T_h - T_c}{R_t} + \frac{1}{2} I^2 R_e \right) \quad (2)$$

The power input to the TEC, P_{TEC} , is equal to difference between Q_h and Q_c . That is,

$$P_{TEC}(I, T_c, T_h) = Q_h - Q_c = N(SI\Delta T + I^2 R_e), \quad (3)$$

where ΔT is equal to $T_h - T_c$. The efficiency of a TEC is characterized by two factors: 1) the maximum heat pumping capability, and 2) the figure-of-merit $Z = S^2/(\rho k)$. Superlattice TECs achieve their superiority over standard TECs in both factors because of their very thin profiles [5].

III. PROPOSED SYSTEM-LEVEL MODELING METHOD

A. Power Model

We need to model the *per-core power consumption* in relation to the processor's operating settings in terms of the DVFS settings, i.e., supply voltage V and frequency f . Different workloads trigger to variations in power consumption even at the same DVFS setting. Thus, we utilize workload-specific information such as number of retired Instructions Per Cycle (IPC). Our model is given by

$$P(f, V, IPC) = w_0 + w_1 f V^2 + (w_2 - w_3 f) IPC^{w_4}, \quad (4)$$

where w_0, \dots, w_4 are weights to be learned from characterization data of the real processor [2].

B. Thermal Model

Our steady-state lumped thermal model is given in Figure 2. Each layer is modeled as follows.

Layer 1: The target multi-core die consists of a two-dimensional grid of cores. If m denotes the total number of cores, then a square grid, for example, will have \sqrt{m} cores per row. We have modeled the lateral heat transfer between adjacent cores through the thermal resistances $R_{(die)i,i-1}$, $R_{(die)i,i+1}$, $R_{(die)i,i+\sqrt{m}}$, and $R_{(die)i,i-\sqrt{m}}$ and the vertical heat transfer between Layer 1 and Layer 2 through $R_{(die)i}$. The values of these resistances depend on the area of the cores, the die thickness, and the thermal conductivity for silicon. Each

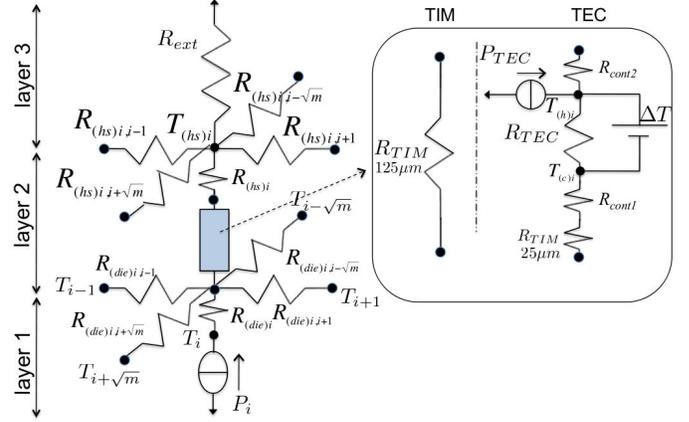


Fig. 2. Steady-state lumped thermal model for a core and its associated TEC.

core dissipates power, P_i , as given by Equation (4) and has a temperature T_i .

Layer 2: In current processors layer 2 is composed of a thermal interface material (TIM) of about $125 \mu\text{m}$ in thickness. When a superlattice is inserted in the volume between the processor and the heat spreader, it takes $100 \mu\text{m}$ in length and $25 \mu\text{m}$ will be the length of the TIM [5]. The TEC is modeled with three elements: its thermal resistance, R_{TEC} , a temperature source $\Delta T_i = T_{(h)i} - T_{(c)i}$ corresponding to the temperature gradient that the TEC i establishes between its plates, and a heat source P_{TEC} . We assume the use of one independent superlattice TEC per core, with an area equal to that of the core. We also considered the presence of contact parasitics which degrades the performance of the TEC. There are thermal contact resistances between the TEC and the TIM (R_{cont1}) and between the TEC and the heat spreader (R_{cont2}) [5]; in the latter case, the thermal resistance of the solder must be also considered.

Layer 3: We modeled the heat transfer through the heat spreader using the vertical thermal resistance $R_{(hs)i}$ and the lateral thermal resistances $R_{(hs)i,i-1}$, $R_{(hs)i,i+1}$, $R_{(hs)i,i+\sqrt{m}}$, and $R_{(hs)i,i-\sqrt{m}}$. The values of these resistances depend on the area of the heat spreader, its thickness, and the thermal conductivity for copper. The temperatures of heat spreader nodes, $T_{(hs)i}$, are the hot-side temperatures for the TECs. We modeled heat transfer through the heat sink and the fan at each node using a lumped thermal resistance R_{ext} . The exact value of this resistance is a function of the size of the heat sink and the fan's speed. When the TECs are engaged, the third layer has to dissipate the sum of heat of both the processor and the TECs.

Given the power consumption of the cores and the currents of the TECs, we solve the system described by the the node equations of the thermal circuit of Figure 2 and Equations (1), (2), (3) to compute the temperatures of all the nodes in the thermal circuit. Due to the nonlinearity of the equations of the circuit, it is necessary to use an iterative numerical optimization method to solve the system. We use MATLAB's `fsolve` function, which is based on the trusted-region algorithm [3].

IV. SYSTEM-LEVEL OPTIMIZATION STRATEGIES

We propose system-level management methods where superlattice TECs, in conjunction with DVFS, are used to mitigate dark silicon problems by reducing hot spot magnitudes, which can be leveraged to either (1) enable large number of cores, or to (2) boost the frequencies of a few cores. The choice between the two options is a function of the workload characteristics and their scalability.

If IPC_i denotes the IPC for the workload executing on core i , then the total throughput of the processor is equal to $\sum_{i=1}^m IPC_i f_i$. Depending on the characteristics of the application, the IPC might have a small dependency on f_i . The throughput equation, $\sum_{i=1}^m IPC_i f_i$, is acceptable when the workload threads running on the cores are heterogeneous with no or little dependency among them. For multi-threaded applications, the IPC of some core i , $IPC_i(h)$, is a function of the number of number of the threads, $h \leq m$, of the workload. Ideally, the total IPC, $\sum_{i=1}^m IPC_i(h)$, should scale linearly with respect to h if the application is perfectly parallelizable; however, in many cases, the total IPC does not scale up, which leads to dark silicon [6], because a portion of the cores cannot improve the performance and thus remains unused.

The goal of our optimization formulation is to maximize the throughput of the processor under a thermal constraint T_{\max} for a given power budget, P_B , for either the TECs or both the TECs and the processor. For heterogeneous workloads, the decision variables for the management system are the DVFS settings (f_i, V_i) of the cores, the currents of the TECs I_i . For multi-threaded workloads, we consider an additional decision variable which is the number of threads, h , for the application.

$$\textbf{Objective:} \max \sum_{i=1}^m IPC_i f_i \text{ or } \max \sum_{i=1}^m IPC_i(h) f_i, \quad (5)$$

such that

$$\forall i : 1 \dots m : T_i \leq T_{\max}, \text{ and} \quad (6)$$

$$\sum_{i=1}^m P_{TEC}(I_i, T_{(c)i}, T_{(h)i}) \leq P_B, \text{ or,} \quad (7)$$

$$\sum_{i=1}^m P(f_i, V_i, IPC_i(h)) + P_{TEC}(I_i, T_{(c)i}, T_{(h)i}) \leq P_B. \quad (8)$$

To find the optimal values for the decision variables, we use an iterative non-linear solver which uses the trusted-region-reflective algorithm [3].

V. EXPERIMENTS

To parameterize our system models, we collected an extensive set of measurement data from an Intel quad-core Core i7 940 processor using the SPEC CPU06 and PARSEC workload. The DVFS settings start at 1.60 GHz; the upper DVFS limit is determined by the maximum allowed temperature of 65 °C. We measured the processor's current consumption using an Agilent 34410A multimeter. We measured the IPC through the performance counters using the `pfmon` package and the temperatures of the cores using the `lmsensors` package.

A. Model Parameterization and Validation

Using the measurements, we identified the weight coefficients of the power model given in Equation (4) through non-linear least-square regression. To verify the model, we ran mixtures of four benchmarks at various DVFS settings; the results from power model deviates with an average of 3.12% from the real measurements.

To parameterize the thermal model, we considered the dimensions of the processor's die from the published layout of the Core i7 and divided the chip's area to four equal stripes, each one corresponding to one core with its portion of the cache. We also assumed standard thermal conductivities for silicon and TIM of $130 \text{ Wm}^{-1}\text{K}^{-1}$ and $1.75 \text{ Wm}^{-1}\text{K}^{-1}$ respectively. To model the thermal resistances related to the heat spreader, we measured its dimensions and considered the thermal conductivity for copper. The external resistance of the sink and fan assembly is equal to $R_{ext}=2.8 \text{ KW}^{-1}$ per core. To verify the thermal models, we considered a TIM for the second layer and compared our model's estimates to the measurements from the thermal sensors. Our results show that the difference in hot spot measurements between the model and sensors is on the average of about 0.8°C. To parameterize the superlattice TEC and the contact resistances R_{cont1} and R_{cont2} , we used the standard parameters published in [5]. We consider four independent TECs, where the area of each TEC is designed to match the area of the underlying core, which leads to $N = 260$ thermocouples per TEC.

B. Results from Optimization Strategies

In the *first experiment*, we demonstrate that using superlattice TECs in conjunction with DVFS provides a better choice over DVFS to meet the thermal constraints despite the additional power required by the TECs. We set the total power budget for both the processor and the TECs to be equal to 73 W. When the TECs are disabled, it is possible to use the power budget to clock the four cores at 2.66 GHz, with a throughput of 12.7 GIPS for the heterogeneous workload set `bwaves - gcc - gobmk - omnetpp`, reaching a hot spot temperature of 60°C. While fixing the total power budget, we varied the thermal constraint from 60°C down to 46°C. Figure 3 shows how the performance (y-axis), expressed as throughput (GIPS), changes versus the temperature (x-axis) for both TECs+DVFS (red solid line) and only-DVFS (blue dotted line). Compared to DVFS alone, TECs allow higher throughput, while still observing the thermal constraint. For example, when the power budget is 73W and $T_{\max}=53^\circ\text{C}$, using TECs give an average improvement of 12% over DVFS alone. We breakdown the total power consumption between the processor and the TECs in Figure 4. The figure shows that when TECs are not used, the processor is not able to use the entire power budget because of arising hot spots. By allocating a portion of the unused power to the TECs to control hot spots, we can use the remaining portion of the unused power to improve performance.

In the *second experiment*, we demonstrate how superlattice TECs can be used to mitigate the first cause leading to

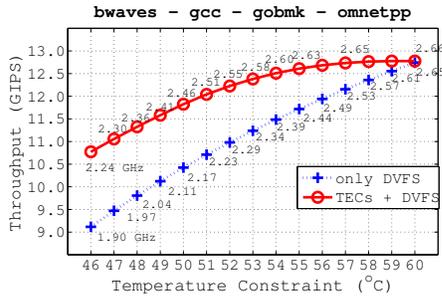


Fig. 3. Maximum possible frequency for a fixed power budget of 73W. Combination **bwaves - gcc - gobmk - omnetpp**.

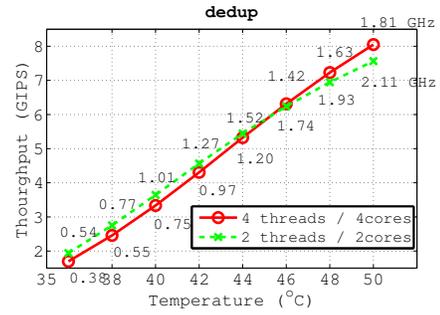


Fig. 6. Results on dedup workload.

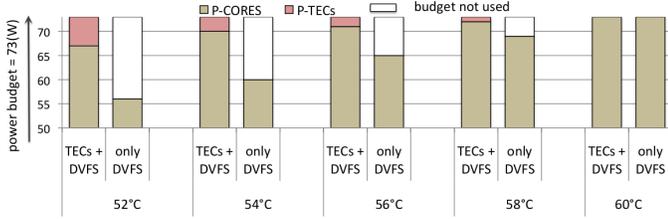


Fig. 4. Allocation of a total budget of 73 W by TECs+DVFS and only DVFS for the workload mix in Figure 3.

dark silicon. We launch four independent instances of the benchmark *hmmmer* from the SPEC CPU 06. We examine for several thermal constraints the number of active cores. Figure 5 gives on the y-axis the number of the dark cores, and the x-axis gives the power of the TECs. When the thermal constraint is 56°C, 3 cores are forced into darkness without engaging the TECs. Allocating 2 W among the TECs enables us to cut down the number of dark cores to 2. This number is further cut to 1 when allocating 4 W. We repeat our experiment for heterogeneous set of SPEC workloads and report the number of dark cores for DVFS alone and for DVFS + TECs in Table I. The results confirm that engaging TECs reduces the number of dark cores.

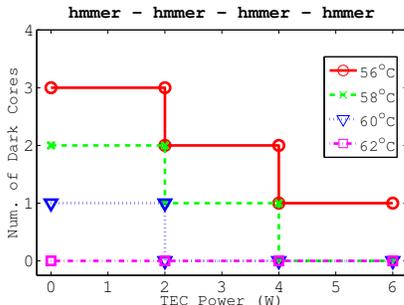


Fig. 5. Mitigation of dark silicon as a function TEC power.

workload mix	# dark cores	
	Only DVFS	DVFS + 4W TECs
tonto - h264ref - xalancbmk - gcc	3	0
h264ref - bzip2 - GemsFDTD - tonto	3	1
gamess - soplex - h264ref - milc	2	0
perlbench - dealll - leslie3d - calculix	3	1
namd - GemsFDTD - bzip2 - sjeng	2	0
mcf - astar - sjeng - gromacs	1	0
cactusADM - zeusmp - gromacs - gamess	2	0
zeusmp - leslie3d - gcc - lbm	3	0

TABLE I

NUMBER OF CORES FOR DVFS AND FOR DVFS + TECs FOR DIFFERENT HETEROGENEOUS WORKLOAD MIXES. THERMAL CONSTRAINT = 55°C.

In the *third* experiment, we consider the second source of dark silicon, when the workload lacks sufficient parallelism. For this case, we focus on the multi-threaded PARSEC benchmarks. The most interesting case is for *dedup*, its IPC per core is 1.99, 1.79, 1.11 for 1, 2, and 4 threads respectively. Figure 6 shows on the x-axis the temperature and on the y-axis the throughput. Once the thermal constraint drops below 44°C, it is better to use two threads, and boost the frequencies of their cores with TECs to outperform the case of four threads. We used a 6 W budget for the TECs.

VI. CONCLUSIONS

In this paper we investigated the use of superlattice-based thermoelectric coolers to mitigate the problems leading to dark silicon in multi-core processors. We developed power and thermal models that are suitable for system-level management and proposed new methods to improve the performance of multi-core processors through a set of optimization methods that leverage the TECs, DVFS, and the number of workload threads. We showed how TECs can be used to mitigate dark silicon problem and improve performance by either enabling more cores or boosting the frequencies of fewer cores.

REFERENCES

- [1] B. Alexandrov, O. Sullivan, S. Kumar, and S. Mukhopadhyay, "Prospects of active cooling with integrated super-lattice based thin-film thermoelectric devices for mitigating hotspot challenges in microprocessors," in *IEEE ASPDAC*, 2012, pp. 633–638.
- [2] A. Bartolini, M. Cacciari, A. Tilli, and L. Benini, "Thermal and energy management of high-performance multicores: Distributed and self-calibrating model-predictive controller," *IEEE Transactions on Parallel and Distributed Systems*, vol. 99, no. PrePrints, 2012.
- [3] M. Celis and R. Dennis, J. and Tapia., "A trust region strategy for nonlinear equality constrained optimization," in *Numerical Optimization*. SIAM, 1985, pp. 71–82.
- [4] P. Chaparro, J. González, Q. Cai, and G. Chrysler, "Dynamic thermal management using thin-film thermoelectric cooling," in *Proceedings of the international symposium on Low power electronics and design*. ACM, 2009, pp. 111–116.
- [5] I. Chowdhury *et al.*, "On-Chip Cooling by Superlattice-based Thin-Film Thermoelectrics," *Nature Nanotechnology*, vol. 4, no. 4, pp. 235–238, 2009.
- [6] H. Esmailzadeh, E. Blem, R. St. Amant, K. Sankaralingam, and D. Burger, "Dark silicon and the end of multicore scaling," in *ACM ISCA*, 2011, pp. 365–376.
- [7] A. N. Nowroz, R. Cochran, and S. Reda, "Thermal Monitoring of Real Processors: Techniques for Sensor Allocation and Full Characterization," in *DAC*, 2010, pp. 56–61.
- [8] Y. Shabany, *Heat Transfer: Thermal Management of Electronics*. CRC Press, 2010.
- [9] O. Sullivan, M. Gupta, S. Mukhopadhyay, and K. S., "Array of thermoelectric coolers for on-chip thermal management," *ASME, Journal of Electronic Packaging*, vol. 134, no. 021005, pp. 1–8, 2012.