Energy Neutral Hybrid Cooling System for High Performance Processors

Luca Rizzon, Maurizio Rossi, Roberto Passerone and Davide Brunelli University of Trento via Sommarive 9, I-38123, Trento, Italy {name.surname}@unitn.it

Abstract-We present the design and testing of a hybrid energy neutral cooling system for data centers' CPUs. The system operates as a passive heat-sink at normal operating conditions, and can provide active cooling when a boost in performance is required (i.e., overclocking) at zero cost by exploiting thermoelectric generators (TEGs) to harvest the energy from the CPU heat dissipation. Server rooms have plenty of wasted heat that can be used to drive low power subsystems that work in symbiosis with servers. Our prototype has been tested on ARM based CPUs, the future core of low power data centers. We firstly evaluated the most suited TEG to supply the conditioning circuit which drives the fan. Then, we defined a governor that allows the active cooling to be exploited when overclocking is necessary and that provides several thermal management policies. Finally we built the proposed system to evaluate the real performance. Experimental results demonstrate its effectiveness in passive and active cooling achieving up to 13 minutes overclocking by recovering the dissipated heat.

I. INTRODUCTION AND STATE OF ART

Any computing device generates heat proportionally to the operating frequency, the supply voltage, and the fabrication technology. High-performance general-purpose microprocessors (CPUs and GPUs) generate a lot of heat. High temperatures, however, may cause malfunctions and micro structural damages that compromise the CPU functionalities. A strategy to prevent overheating is the adoption of Dynamic Voltage and Frequency Scaling (DVFS). DVFS dynamically reduces the clock frequency and/or the supply voltage in order to decrease the temperature at the cost of reduced computing performance. This technique is commonly used to manage temperature in portable devices, where packaging constraints make it impossible to use active cooling [6], [2]. Besides HW and SW thermal management techniques, a microprocessor can be cooled down adopting passive, or active heat dissipation devices. Passive coolers are physical devices, such as heatsinks, that guarantee a proper thermal exchange between the hot microprocessor package and fresh air. Fans, Peltier cells, and other advanced coolers are considered active because they require electric power to remove wasted heat.

Solutions to reduce heat include the use of low power ARM processors. These processors are characterized by lower power consumption and lower thermal dissipation compared with previous generation of personal computer, while still providing high computing performance. Moreover, studies forecasts a migration from traditional architectures toward ARM based devices also in the data centers market, since ARM devices

978-1-4799-6177-1/14/\$31.00 © 2014 IEEE

offer a good trade-off between power consumption and computing performance [7]. Generally, data centers facilities are always turned on, and are sources of high thermal waste. In fact, only 0.5% of total fossil fuel power is spent in useful computation, while more than 65% is generally used for cooling the surrounding environment or lost as heat [3].

In this work we describe an energy neutral cooling device that combines passive with active cooling (hybrid). Active cooling via forced convection is performed using a fan that spins thank to the electrical energy harvested from the heat dissipated by the computing device itself, as depicted in Fig. 1. The proposed system consists of a Thermal Energy Recovery System (TERS) and a fan on top of a heat sink. The TERS is made of a thermoelectric module, a conditioning circuit and two supercapacitors to store the harvested energy. The prototype has been realized and proven to be effective to guarantee thermal dissipation in passive mode, as well as to provide boost in cooling capability in active mode that permits a performance boost to be performed periodically.

In the literature similar designs have been proposed [12]. Energy Harvesting is nowadays a well-known technology [1], and converting thermal energy from electronic computing devices into electric energy has been already demonstrated [8]. However, recovered energy is too small to be used to extend the system battery lifetime [10]. Nevertheless, recent articles [11], [9] claim that the energy can be used to supply cooling devices or context aware computing. Moreover, the implementation increases CPU temperature up to 25°C when the TEG is applied on top of multicore desktop microprocessors, causing the CPU to work near its allowed temperature limit [13].

We demonstrate that energy can be used to supply an active cooler. In normal workload conditions, the system proposed in this paper provides to the microprocessor passive cooling. In this condition, the thermal waste is converted into electrical energy and stored in supercapacitors. The harvested energy is used to supply the active cooling system. Active cooling capability is useful to manage a burst of workload, during which the system runs at its maximum clock frequency.

This article is organized as follows, in Sec. II we describe the HW components used, their characteristics and performance that allowed us to design the prototype system presented in Sec. III. Sec. IV illustrates pros and cons through experimental results obtained with the prototype hybrid cooling system. Finally Sec. V concludes the work.



Fig. 1. Schematic view of the target application scenario. The energy wasted as heat can be collected and used for active cooling with fan.



Fig. 2. Thermoharvester schematic made of 2 Peltier cells, spreaders and sink in stacked fashion. Thermal grease guarantees maximum thermal contact.

II. THERMO-ELECTRIC HARVESTING

Thermoelectric generators (TEG) are devices that convert a thermal gradient ΔT into electrical energy. TEGs output power increases almost linearly with the increase of ΔT but the device performance and lifetime are guaranteed only if the absolute hot-side temperature is lower than 80°C [5]. This upper bound corresponds to the same threshold below which the ARM CPUs work in safe condition. Conversely, a kernel governor can implement thermal management strategies as soon as this threshold is crossed.

There exist several commercial devices that allow thermal energy to be harvested and converted into electrical energy. We have compared three different devices to select the one that offers the best features for our aim. The devices we chose as TEGs are: (a) *Nextreme* eTEG HV56 Thermoelectric Power Generator without its output power regulator; (b) *Micropelt* TE-CORE7 TGP-751 ThermoHarvesting

Power Module with a 33 mm heat sink; (c) *Peltier-Cell* PE1-12706AC 40×40 mm squared cells. The last ones are designed and intended for cooling applications, therefore it has good performance as a cooler, and it exhibits mediocre performance as a thermoelectric generator. Moreover, we use them only as



Fig. 3. Output Power, Nextreme on Arndale.

harvesters, because the current needed to operate as coolers is very high (in the Ampere range) compared with the scavenged current (see results later).

Peltier cells were arranged in two different configurations, (i) single cell and (ii) double cell connected in series and arranged in stacked fashion, as depicted in Fig. 2. The thermal contact between the ARM CPU and the TEG is optimized by means of an aluminum layer - 2 mm thick - that acts as a spreader, while the thermal exchange with the environment is maximized by the heat sink on top of the stacked harvester. In configuration (ii) we add an extra thin aluminum layer (0.2 mm)in between the Peltier cells to ease the flow of the thermal energy.

The target high performance computing board is a Samsung Arndale equipped with an Exynos, 1.7 GHz dual-core ARM Cortex A15 processor, running Linux kernel v. 3.10. We implemented different tasks in sequence to evaluate its performance. In particular the tasks were selected to achieve a range of time lengths and CPU loads (with fixed maximum clock frequency) that are directly related with CPU temperature: (a) video encoding using ffmpeg (GPL software) with four threads, that converts a movie; (b) multi-thread application that performs millions of algebraic and trigonometric computations of floating point numbers using X threads (called busywork-X in pictures); (c) kernel operations in this task a Linux kernel is uncompressed, compiled and then cleaned, finally the folder is removed. We measured the impedance matched power - as the product of voltage and current over a matched load with a 1 s period - to compare the performance of the TEGs and select the most suited one to power the cooling system.

Detailed results of the output profile obtained by running the tasks are reported in Fig. 3 to Fig. 6. These pictures present the output impedance matched power of each TEG exposed to the ΔT produced by the heat dissipated by the CPU. The thermal gradient in turn is strictly related with the CPU load. The air conditioning system of the lab kept the room temperature almost constant in the range 20° C to 25° C.

The most interesting insight is related to the CPU temperature. Notice that the commercial TEGs cannot dissipate completely the underlying heat when the CPU is pushed to the maximum performance. In case of Nextreme, the ΔT rises up to 52°C while for Micropelt it reaches 60°C. Those values correspond to the critical CPU temperature above 80°C. Consequently the kernel's governor scales down the clock



Fig. 4. Output Power, Micropelt on Arndale.



Fig. 5. Output Power, Peltier cell on Arndale.

frequency (f_{CLK}) resulting in the "flickering" evolution during busywork-100 clearly evident in Fig. 4 and less remarkably in Fig. 3. Contrarily, this behavior does not occur when using the stacked Peltier cells with a custom heat-sink. Considering the detailed results of Tab. II, we can see that Peltier cells reduce the thermal gradient by more than 20°C (single cell) and 5°C in case of double cells with respect to Nextreme TEG at the maximum operating point (busywork-100). Also, from the pictures we can see that TEGs exhibit a good reactivity to fast temperature spikes, resulting in sharp peaks at the beginning of each task; only the Nextreme one produces smoothed power profiles indicating a worst thermal efficiency. Finally, comparing Fig. 3 to 6, we can also notice that the custom TEG + heat-sink solution demonstrates the best thermal dissipation since the maximum registered ΔT is $\approx 50^{\circ}$ C (double Peltier at maximum CPU load) which corresponds to $\approx 75^{\circ}$ C.

The two Peltier cells in series almost double the output power of the single cell: the average matched power raises from 2.6 to 3.6 mW at the maximum CPU load, as reported in Tab. III. Generally, with tasks longer than 30 s the stacked Peltier produces the highest output, while with very short tasks (the last three of each benchmark) the Micropelt TEG overcomes the others: 1.593 mW in case of busywork-10 against 0.837 mW of the double Peltier.

Results in terms of energy and execution time are summarized in Tab. I. The stacked architecture can harvest the highest amount of energy in the least time (26 J in 4^h 30^m with double Peltier cell vs. 24 J in 5^h 45^m with Nextreme and 20 J in 5^h 20^m with Micropelt). Micropelt and Nextreme



Fig. 6. Output Power of two Peltier cells.

TABLE I.	TOTAL BENCHMARK EXECUTION TIME AND ENERGY
	HARVESTED.

TEG	Energy [J]	Time [min]
Micropelt	19.9	318.9
Nextreme	22.9	343.7
1-Peltier	18.9	255.0
2-Peltier	26.8	258.9

TEGs can not dissipate the heat produced by the Arndale's CPU, and - as a consequence - those are not able to fully exploit the high variations in ΔT . Moreover, this results also in a longer execution time of the benchmark because the DVFS kernel module slows down the CPU to reduce its temperature. We selected the series of two Peltier cells arranged in stacked fashion as starting point to design and test the hybrid energy neutral cooling system described in the following section.

III. PROTOTYPE COOLING SYSTEM.

In the previous section, we described the characteristics of an example high performance ARM CPU running at maximum speed. Generally, data centers machines work at $f_{CLK} < f_{CLK}^{MAX}$ taking advantage of the redundancy. However, it is sometimes necessary to run at maximum performance, even overclocking the machine to speed up computations and/or serve concurrent processes. In these cases, the thermal management is crucial to guarantee the safety of the computing units. To address this scenario, the energy neutral hybrid cooling system works in symbiosis with the CPU.

The target system has been described in Sec. I. We implemented it to execute tests introducing an external circuit for storage and conditioning, and the governor that runs on



Fig. 7. Experimental setup block diagram.

TABLE II. Average thermal gradient ΔT on the TEGs with respect to the tasks.

TEG	ΔT [°C] vs. Task							
	idle	ffmpeg	busywork-100	tar	make	clean	rm	busywork-10
Micropelt	15.0	34.9	57.5	22.9	45.4	22.6	21.5	35.0
Nextreme	15.0	29.9	51.6	20.4	43.3	19.5	16.9	26.4
1-Peltier	6.5	13.9	21.9	9.0	17.0	7.9	9.2	11.3
2-Peltier	16.0	30.3	45.2	20.9	32.9	19.8	18.3	23.3

TABLE III. AVERAGE IMPEDANCE MATCHED POWER PER TEG WITH RESPECT TO THE TASKS.

TEG	Impedance Matched Power [mW] vs. Task							
	idle	ffmpeg	busywork-100	tar	make	clean	rm	busywork-10
Micropelt	0.320	0.841	2.278	0.352	1.547	0.342	0.410	1.593
Nextreme	0.198	0.779	2.515	0.318	1.709	0.275	0.348	1.318
1-Peltier	0.378	1.026	2.694	0.349	2.373	0.242	0.252	0.377
2-Peltier	0.475	1.573	3.585	0.488	2.915	0.381	0.403	0.837

an external PC. The whole experimental setup is depicted in Fig. 7. The storage and conditioning circuit collects the electrical energy provided by the double Peltier TEG and stores it into supercapacitors. Fig. 7 depicts the details of this circuit which has been made with commercial components. The fan is a 40x40x5 mm rated at 5 V and 250 mA. By experimental evaluation, we measured the minimum voltage (4.4 V) required to switch it on and reach a speed (we used a two wire fan with no feedback on actual rpm) sufficient to cool down the heat-sink within the safety threshold (discussed later). Two 50 F 2.5 V supercapacitors were used as storage unit, because of the high energy required to switch on the fan with enough speed. Supercapacitors are used in mixed mode, they are charged in parallel to speed up the process, then switched in series to reach the 4.4 V required by the fan. Fig. 8 demonstrates the effectiveness of this choice in terms of voltage stored in the two configuration with respect to time (measured using supercapacitor model [4]). The serial configuration takes $5 \times$ more time to collect the same amount of energy with respect to the parallel one (≈ 600 s versus \approx 130 s). Considering the average energy collected by the TEG, this is still an improvement in speed of recharge even if it has been measured in the order of ten hours (see Sec. IV). Supercapacitors are charged up until the voltage across them reaches the threshold TH_{cap} . The value of TH_{cap} has been set equal to 2.2 V because capacitors are charged in parallel and the series equivalent 25 F provides enough current to switch on the fan (thanks also to the low internal resistance of the supercapacitors).

The comparator is used to check the storage charge status, its thresholds are V_{MAX} equal to 2.2 V and V_{MIN} equal to 2.0 V. The high level allows the fan to be switched on when enough charge is stored in the reservoir, while the low one is not used since the fan's switch off is handled by the governor. Finally, the circuit includes - even if those are not shown - the switches to commutate the storage elements and to turn on the fan by direct connection; both are realized with relays.

The governor, which is summarized in the flowchart of Fig. 9, extends the kernel's DVFS module. It has been tested in Matlab, and it uses serial communication to interface with the Arndale board. The measurement instrumentation is connected to the PC running the governor: multimeters to measure fan's current and comparator status, and a programmable voltage supplier to drive the relays. We defined two working scenario to characterize the cooling device response in passive and active modalities: (a) **Standard**, the CPU runs with intermediate f_{CLK} , here **Passive Cooling** safeguards the system; (b)



Fig. 8. Voltage stored in two 50 F supercap in time, parallel versus series configurations.



Fig. 9. The governor's flowchart.

Overclock, the CPU runs at maximum speed and load, here **Active Cooling** extends the time spent with overclock enabled still keeping the CPU within safety thresholds. The modalities are managed by the governor, which periodically checks the comparator to know the storage's state of charge. If the stored energy is above TH_{cap} we can afford an overclocking phase, otherwise we must wait for the TEG to collect further energy. An improvement will be to let the governor check the queue of active processes and, based on a threshold (TH_{stack}) , decide whether or not switch the speed to overclock. To conduct the tests presented in the next section we chose to activate the overclocking immediately after the change in the comparator status.

The fan is not switched on simultaneously with the overclock start. This choice allows us to extend the overclocking



Fig. 10. Comparison of CPU temperature rise with 100% CPU load @ 1.5 GHz. Continuous line depicts natural heat dissipation, above 85°C the kernel governor scales down the f_{CLK} . The heat-sink increases the performance and Peltier cells ease the thermal flow (2°C gain).

total time and the energy collected by the TEG because (i) the temperature rises up slowly thanks to the presence of the TEG plus heat-sink, and (ii) the power harvested increases with ΔT . While in overclock mode, the governor continues to check the CPU temperature periodically (delay) until it matches the (TH_{CPU}) threshold. At this stage the fan is switched on (for programmable time amount) and the flag is set to true. After the active cooling phase, the overclock can continue because of the decrease in temperature. The second time the TH_{CPU} is exceeded, the flag is true, hence the governor slows down the f_{CLK} to the normal operating mode, since not enough energy remains in the storage unit to supply the active cooling. In the following section we enter in the detail of the current implementation before presenting the performance of the proposed system.

IV. EXPERIMENTAL RESULTS

In this section we present the measured performance of the energy neutral cooling system. We conducted the tests by fixing threshold and modalities, introduced in the description of the prototype, as $TH_{CPU} = 68^{\circ}$ C while the CPU workload modalities are: (a) Standard, the CPU runs with 100% load @ 1.5 GHz. (b) Overclock, the CPU runs with 100% load @ 1.7 GHz. We set a limit to block the simultation if the CPU temperature exceeds 70°C. This value of temperature represents a maximum working point to avoid damages for both the CPU and the TEGs, as described in Sec. II. The workload instead is achieved by executing the busywork task continuously. Even if a continuous fixed workload is not realistic to model the usage of data-centers and high performance computing machines in general, this choice is justified by the fact that we wanted to evaluate both the performance in terms of energy harvested from the heat and the passive heat dissipation in very harsh conditions. Data-center modeling is out of the scope of our project. It would not be possible to test all the possible workload scenarios and there is no benchmark to evaluate TEGs performance with computing units, so we designed the above to be simple and easily reproducible on any other platform. Obviously, using the CPU with such a high workload speeds up the supercapacitors recharging phase which is the limit of the current implementation.

i. Passive Cooling

We evaluated the thermal dissipation performance of the



Fig. 11. Long term stability of the passive cooling with 100% CPU load @ 1.5 GHz. Time required to recharge the supercap.

heat-sink plus double Peltier TEG in comparison with natural heat dissipation and the heat-sink, in the standard workload case. Results are depicted in Fig. 10. The continuous line depicts the temperature profile of the Arndale's CPU without heat-sink, as it is sold. The dashed line depicts the resulting temperature evolution with the custom heat-sink, while the fined dashed one shows the performance of the proposed harvesting system. In the first case, the CPU temperature rises up to the safety threshold of 85°C (imposed by the Linux kernel's DVFS) within 5 minutes. In the two other cases the DVFS does not need to intervene, since the CPU stays within this threshold. Notice the proposed harvesting stacked structure provides 2°C gain with respect to the other solutions. Thanks to the wider side surface exposed to the air, and the presence of more spreader in between the two stages, the proposed harvesting stacked structure provides 2°C gain with respect to the other solutions.

Fig. 11 depicts the time interval required by the prototype circuit to recharge the storage supercapacitors. These results are part of the active cooling performance evaluation presented later, but it is meaningful to underline the ≈ 10 hours required. This picture also presents the long term stability of the thermal performance just described, since in between the two spikes (two active cooling instants) the CPU temperature is basically constant (the top curve that oscillates between 62° C and 63° C). In this picture the current is the TEGs output that charges the storage, f_{CLK} is the feedback on active modality (standard vs. overclock) and the lower stable dotted line represents the environmental temperature.

ii. Active Cooling

The active cooling tests required very long time as introduced above, so we let the system run for several days to extract the following results. We also evaluated different interval delays to select the best suited with the target hardware architecture. We evaluated from 2 to 30 s, but the most interesting are with 10 and 30 s cases. The 30 seconds represent the limit we imposed because of the recharging. An example result in this case is depicted in Fig. 12. Here the CPU temperature increases twice above to the threshold $TH_{CPU} = 68^{\circ}$ C, the first occurrence triggers the drop in the charge current (which means the stop of the recharge to power the fan); the second time it triggers the end of the overclock phase.

The 10 seconds active time instead is interesting because



Fig. 12. Active cooling, fan activity of 30 s allows 13 min overclocking @ 22.5° C.



Fig. 13. Active cooling, fan activity of 10 s allows 13 min overclocking @ 20°C.

it underlines the very impacting effect of the environmental temperature on the performance. In this case we obtained the same overclocking length in time of 13 min we get in the 30 s case with only 2.5°C difference in the environment. With temperature same as above the 10 s activity resulted in few minutes of overclocking (picture omitted for space reasons). In this fan activity configuration otherwise we expect to achieve three overclocking phases in the 10 hours reference time, providing more flexibility to the final system.

iii. Final Observations

Performance and stability of the system depend mainly on the environmental temperature as it has been shown, not only on the workload. Other impairing factors are related with the tolerance of the prototype circuit where wirings, soldering, and the use of breadboards affect the real thresholds. Also the laboratory's HVAC system and its characteristics (position of the HVAC inlets, direct sunlight, isolation, windows, etc.) did not allow us to have uniform conditions during the long experiments (several day each). Despite these unavoidable circumstances, to summarize the results we can state that the proposed energy neutral cooling system can achieve up to 13 min of overclocking every ≈ 10 hours. In case of very stable environmental temperature around 20°C, we demonstrate even better performance, but further investigation is required. What we want to evaluate, in particular, is the possibility to split the use of harvested energy in more dynamic ways: depending on the number of stacked processes in the execution queue and the energy reserve level.

V. CONCLUSION

In this work we describe an energy neutral hybrid cooling system that combines passive, and active cooling in a single device. Active cooling via forced convection is performed using a fan that spins thank to the electrical energy harvested from the heat dissipated by the computing device itself. Passive cooling is provided by the energy harvesting device composed of two Peltier cells connected in series and a heat-sink, arranged in a stacked fashion. We characterized the response of several commercial and custom TEGs to select the most suitable to use with high-performance computing units based on ARM processors. A demonstrator has been realized, and proven to be effective to guarantee thermal dissipation in passive mode, as well as provide boost in cooling capability in active mode, that allows us to control the CPU temperature during overclocking. The prototype energy neutral cooling system can achieve up to 13 min of overclocking every ≈ 10 hours.

These results pave the way to a new kind of cooling systems for future high-performance processors. We plan to improve the prototype both from technological (integration of the conditioning circuit into an optimized board, new tailored supercapacitors) and functional (governor as kernel task) points of view.

ACKNOWLEDGMENTS

The research contribution presented in this paper has been supported by project GreenDataNet (grant n.609000) funded by the EU 7th Framework Programme.

REFERENCES

- D. Carli, D. Brunelli, L. Benini, and M. Ruggeri. An effective multi-source energy harvester for low power applications. In *Design, Automation Test in Europe Conference Exhibition (DATE), 2011*, pages 1–6, March 2011.
- [2] K. Choi, R. Soma, and M. Pedram. Dynamic voltage and frequency scaling based on workload decomposition. In *Low Power Electronics* and Design, 2004. ISLPED '04. Proceedings of the 2004 International Symposium on, pages 174–179, 2004.
- [3] Data Centre Specialist Group. Meeting the energy efficiency and financial challenges in it. Technical report, BCS, 2007.
- [4] Devillers, N. et al. Review of characterization methods for supercapacitor modelling. *Journal of Power Sources*, 246:596–608, 2014.
- [5] Laird Technologies. Thermoelectric handbook.
- [6] Murali, S. et al. Temperature control of high-performance multi-core platforms using convex optimization. In *Proceedings of the Conference* on Design, Automation and Test in Europe, DATE '08, pages 110–115. ACM, 2008.
- [7] Rick Merritt. Dell, IBM give thumbs up to ARM Servers. In *EE Times*, 2010.
- [8] L. Rizzon, M. Rossi, R. Passerone, and D. Brunelli. Wireless sensor networks for environmental monitoring powered by microprocessors heat dissipation. In *Proceedings of the 1st International Workshop on Energy Neutral Sensing Systems, ENSSys '13.* ACM, 2013.
- [9] M. Rossi, L. Rizzon, M. Fait, R. Passerone, and D. Brunelli. Energy neutral wireless sensing for server farms monitoring. *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, PP(99):1– 11, 2014.
- [10] G. Solbrekken, K. Yazawa, and A. Bar-Cohen. Experimental demonstration of thermal management using thermoelectric generation. In *Thermal and Thermomechanical Phenomena in Electronic Systems*, 2004. ITHERM '04. The Ninth Intersociety Conference on, pages 284– 290 Vol.1, 2004.
- [11] G. Solbrekken, K. Yazawa, and A. Bar-Cohen. Thermal management of portable electronic equipment using thermoelectric energy conversion. In *ITHERM '04. The Ninth Intersociety Conf. on*, pages 276–283 Vol.1, June 2004.
- [12] E. Suski. Method and apparatus for recovering power from semiconductor circuit using thermoelectric device, May 30 1995. US Patent 5,419,780.
- [13] C.-J. Wu. Architectural thermal energy harvesting opportunities for sustainable computing. *Computer Architecture Letters*, PP(99), 2013.