RETROSPECTIVE: Temperature-Aware Microarchitecture

Kevin Skadron*, Mircea R. Stan*, W. Huang†, S. Velusamy‡, Karthik Sankaranarayanan§, David Tarjan¶

*University of Virginia, †ZEKU Technologies, ‡Google,§Intel, ¶NVIDIA

I. CONTEXT

This paper was published at ISCA 2003. At the time, design of the thermal solution was typically done after chip design by a packaging team that used worst-case post-synthesis power numbers and ensuring that temperatures never rose above specified constraints. The thermal model for the chip was typically steady-state and with no spatial detail (i.e., single node). The architecture community was just beginning to look at temporal and spatial variations, accounting for workload phases and differing power dissipation of various units. For example, last-level caches typically have very low power density.

Our work was motivated by emerging work on thermal modeling and management [1], [9], which—motivated by observations about the thermal design for the Intel Pentium 4 [5]—argued for designing processors for an *expected* range of thermal operation, with worst-case behavior handled by dynamic throttling instead of costly worst-case thermal packaging. But prior work approximated runtime temperature using a simple moving average of recent power dissipation. The attention to power in the architecture community, and the advent of micro-architectural power modeling, were themselves fairly recent developments.

We knew that thermal effects exhibit a non-linear response due to *thermal capacitance* and that there would be some thermal coupling among microarchitectural units, so simple, linear averages were not sufficient. At the same time, we also recognized that runtime thermal throttling should target local hotspots and would benefit from feedback control theory, and in fact we had explored the use of feedback control for thermal management in an earlier HPCA 2002 paper [22], which introduced a cruder version of Hotspot. This early version of HotSpot leveraged a well-known duality between electrical and thermal resistance and capacitance, allowing use of standard circuit solvers. However, the HPCA 2002 model did not adequately account for the thermal packaging and lateral heat transfer among microarchitectural blocks.

II. INSIGHTS

Compared to our original HPCA 2002 version, the ISCA 2003 paper developed a more refined thermal model, simplified the feedback control, addressed thermal sensor placement and precision, and proposed some new mechanisms for thermal management, albeit in the context of the single-core processors of the era. A few key insights were:

- There is a need for thermal models that explore both spatial and temporal thermal variations at the granularity of microarchitectural blocks of interest, and that allowed pre-RTL design space exploration. This required the thermal tool to be compatible with pre-RTL computer architecture infrastructure and for the thermal model to adapt automatically based on the microarchitectural configuration (in combination with area models). The possibility of lateral heat transfer among units also required floorplan information, which we originally approximated with hand measurements of published floorplans and simplistic layout. We later developed tools to automate floorplan generation [4], [19].
- Modeling leakage power is essential, because leakage increases with temperature, creating a feedback loop.
- The physical packaging of the die and the role of active cooling (e.g., fans) were essential for accurate thermal modeling. We believe that the most important contribution of HotSpot, and the reason it has found widespread use, is that HotSpot captures enough physical detail so that, with the physical properties of these materials (dimensions and material), and an equivalent thermal resistance to account for the heat transfer to the air, we could derive a complete thermal model *without the need for empirical fitting factors*. Validation against a commercially available test chip [11] showed the importance of the thermal interface material, and further modeling refinements (grid model, high-aspect-ratio blocks, etc.) reached maturity in v. 5.02.
- We explored a variety of microarchitectural thermal throttling mechanisms, including simple frequency scaling, dynamic voltage and frequency scaling (DVFS), fetch throttling, activity migration, and global clock gating, and found that the lowest-overhead approach was to reduce the frequency, because this could be achieved in much smaller increments than other techniques, and with low overhead. Throttling fetch was also effective. In later work, we realized that the choice of technique could be based on the severity of the thermal stress [21].
- For the feedback control, we found a PI controller worked well. The integral component significantly improved convergence, but integral windup needed to be addressed to maintain responsiveness [22].
- The placement and precision of the temperature sensors were essential. The greater the imprecision, the lower the temperature setpoint (i.e., guardband) needed to be

for safe operation, and the more frequently thermal management was engaged, with consequent performance overhead. In this paper, we assumed one thermal sensor per architectural block. We later explored quantity and placement of sensors in a more rigorous fashion [13].

III. Evolution of HotSpot

HotSpot went through several iterations that improved fidelity, accuracy and performance and added features.

- HotSpot 1.0: Original thermal model in ISCA 2003.
- HotSpot 2.0 [11]: Addition of a grid thermal model, validation against physical test chips, added details for accuracy. This version also added support for 3D chip stacks and automated floorplanning [19].
- HotSpot 3.0: Better solvers, added details for accuracy (e.g. die edges).
- HotSpot 4.0 [10]: Better solvers, more accurate models, accounting for blocks with high aspect ratios.
- HotSpot 5.0: Configurable model for the thermal package, more detailed thermal model (especially the secondary heat-transfer path to the circuit board, which accounted for about 10% of the heat transfer).
- HotSpot 6.0: Better solver (SuperLU), further steps towards 3DIC with better support for multiple layers, incorporating work from Ayse Coskun's group [16].
- HotSpot 7.0 [6]: Support for microchannel cooling, informed by many sources but especially 3D-ICE [23].

In addition, we showed the importance of process variations [12], and migrated our modeling framework and architecture work to multicore processors, e.g. [8] [14] and GPUs [20], demonstrated potential thermal denial-of-service attacks [3], and used HotSpot to explore dynamic lifetime-reliability management, e.g. [15].

IV. Thoughts on HotSpot Today

We were fortunate to introduce HotSpot just when concerns about power were crystallizing into several distinct thrusts with different modeling and management criteria: thermal management (requiring HotSpot-style modeling); energy efficiency (where optimizing average power is sufficient, since $E = P_{avg} * t$; and power delivery (where instantaneous changes in power can cause voltage droop, depending on both the power delivery network and short-term microarchitectural behavior). HotSpot has been widely used in academia, and we are thrilled that it enabled researchers and technologists to explore their creative ideas. Notable examples include thermally-aware task scheduling [2]; computational sprinting [17]; limiting average power, allowing for brief thermal overshoots to exploit thermal capacitance (Running Average Power Limit or RAPL) [18]; extensions to server- and rack-level modeling and management [7], and many more.

Research continues in thermal modeling and management for ICs and computer systems, creating many further opportunities for research from VLSI to rack-scale systems and software optimizations at all levels of the stack.

References

[1] D. Brooks and M. Martonosi, "Dynamic thermal management for high-performance microprocessors," in *Int'l Symp. on High Performance Computer Architecture (ISCA)*, Jan. 2001.

[2] J. Choi, C.-Y. Cher, H. Franke, H. Hamann, A. Weger, and P. Bose, "Thermal-aware task scheduling at the system software level," in *Int'l Symp. on Low Power Electronics and Design (ISLPED)*, Aug. 2007.

[3] P. Dadvar and K. Skadron, "Potential thermal security risks," in *Int'l Semiconductor Thermal Measurement, Modeling, and Management Symp. (SEMI-THERM)*, Mar. 2005.

[4] G. Faust, R. Zhang, K. Skadron, M. Stan, and B. Meyer, "ArchFP: Rapid prototyping of pre-RTL floorplans," in *Int'l Conf. on Very Large Scale Integration (VLSI-SoC)*, Oct. 2012.

[5] S. Gunther, F. Binns, D. M. Carmean, and J. C. Hall, "Managing the impact of increasing microprocessor power consumption," in *Intel Technology J.*, Q1 2001.

[6] J. Han, R. E. West, K. Skadron, and M. R. Stan, "Thermal simulation of processing-in-memory devices using HotSpot 7.0." in *Int'l Workshop on Thermal Investigations of ICs (THERMINIC)*, Sep. 2021.

[7] T. Heath, A. P. Centeno, P. George, L. Ramos, Y. Jaluria, and R. Bianchini, "Mercury and Freon: Temperature emulation and management for server systems," in *Int'l Conf. on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, Oct 2006.

[8] W. Huang, K. Rajamani, M. R. Stan, and K. Skadron, "Scaling with design constraints — predicting the future of big chips," *IEEE Micro*, 31(4), July/Aug 2011.

[9] W. Huang, J. Renau, S.-M. Yoo, and J. Torellas, "A framework for dynamic energy efficiency and temperature management," in *Int'l Symp. on Microarchitecture (MICRO)*, Dec. 2000.

[10] W. Huang, K. Sankaranarayanan, R. J. Ribando, K. Skadron, and M. R. Stan, "Accurate, pre-RTL temperature-aware processor design using a parameterized, geometric thermal model," *IEEE Trans. on Computers*, 57(9), Sep. 2008.

[11] W. Huang, M. R. Stan, K. Skadron, S. Ghosh, K. Sankaranarayanan, and S. Velusamy, "Compact thermal modeling for temperature-aware design," in *Design Automation Conf. (DAC)*, Jun. 2004.

[12] E. Humenay, D. Tarjan, and K. Skadron, "The impact of systematic process variations on symmetrical performance in chip multi-processors," in *Conf. on Design, Automation, and Test in Europe (DATE)*, Apr. 2007.

[13] K.-J. Lee, K. Skadron, and W. Huang, "Analytical model for sensor placement on microprocessors," in *Int'l Conf. on Computer Design (ICCD)*, Oct. 2005.

[14] Y. Li, B. C. Lee, D. Brooks, Z. Hu, and K. Skadron, "CMP design space exploration subject to physical constraints," in *Int'l Symp. on High-Performance Computer Architecture (HPCA)*, Feb. 2006.

[15] Z. Lu, J. Lach, M. R. Stan, and K. Skadron, "Temperature-aware modeling and banking of IC lifetime reliability," *IEEE Micro*, 25(6), Nov./Dec. 2005.

[16] J. Meng, K. Kawakami, and A. K. Coskun, "Optimizing energy efficiency of 3-D multicore systems with stacked DRAM under power and thermal constraints," in *Design Automation Conf. (DAC)*, June 2012.

[17] A. Raghavan, Y. Luo, A. Chandawalla, M. Papaefthymiou, K. P. Pipe, T. F. Wenisch, and M. M. K. Martin, "Computational sprinting," in *Int'l Symp. on High-Performance Comp Architecture (HPCA)*, Feb. 2012.

[18] E. Rotem, A. Naveh, D. Rajwan, A. Ananthakrishnan, and E. Weissman, "Power-management architecture of the Intel microarchitecture code-named Sandy Bridge," *IEEE Micro*, 32(2), Mar.-Apr. 2012.

[19] K. Sankaranarayanan, S. Velusamy, M. R. Stan, and K. Skadron, "A case for thermal-aware floorplanning at the microarchitectural level," *Journal of Instruction-Level Parallelism (JILP)*, 2005, vol. 7.

[20] J. W. Sheaffer, K. Skadron, and D. P. Luebke, "Studying thermal management for graphics-processor architectures," in *Int'l Symp. on Performance Analysis of Systems and Software (ISPASS)*, Mar. 2005.

[21] K. Skadron, "Hybrid architectural dynamic thermal management," in *Design, Automation and Test in Europe Conf. (DATE)*, Feb. 2004.

[22] K. Skadron, T. Abdelzaher, and M. R. Stan, "Control-theoretic techniques and thermal-RC modeling for accurate and localized dynamic thermal management," in *Int'l Symp. on High-Performance Computer Architecture (HPCA)*, Feb. 2002, pp. 17–28.

[23] A. Sridhar, A. Vincenzi, M. Ruggiero, T. Brunschwiler, and D. Atienza, "3D-ICE: Fast compact transient thermal modeling for 3D ICs with inter-tier liquid cooling," in *Int'l Conf. on Computer-Aided Design (ICCAD)*, 2010.