The Need for a Full-Chip and Package Thermal Model for Thermally Optimized IC Designs

Wei Huang†, Eric Humenay‡, Kevin Skadron‡, and Mircea R. Stan†
† Charles L. Brown ECE Department, University of Virginia, Charlottesville, VA 22904, USA
‡ CS Department, University of Virginia, Charlottesville, VA 22904, USA
{whuang, humenay, mircea}@virginia.edu, skadron@cs.virginia.edu

ABSTRACT

Modeling and analyzing detailed die temperature with a full-chip thermal model at early design stages is important to discover and avoid potential thermal hazards. However, omitting important aspects of *package* details in a thermal model can result in significant temperature estimation errors. In this paper, we discuss the applications of an existing compact thermal model that models both *die* and *package* temperature details. As an example, a thermally self-consistent leakage power calculation of a POWER4-like microprocessor design is presented. We then demonstrate the importance of including detailed package information in the thermal model by several examples considering the impact of thermal interface material (TIM), which glues the die to the heat spreader. The fact that detailed package information is needed to build an accurate compact thermal model implies a design flow, in which the chip- and package-level compact thermal model acts as a convenient medium for more productive collaborations among circuit designers, computer architects and package designers, leading to early and efficient evaluations of different design tradeoffs for an optimal design from a thermal point of view.

Categories and Subject Descriptors:
B.7.2 [Hardware]: Design Aids
J.6 [Computer-Aided Engineering]: Computer-Aided Design.
General Terms: Design, Reliability.
Keywords: thermal model, package, temperature-aware design, leakage.

1. INTRODUCTION

The continued scaling of CMOS IC technology [1] together with the resultant ever-increasing power density and operating temperature poses significant challenges to designers at the circuit, architecture and package levels. Elevated operating temperature greatly increases the leakage power as IC designs move into the deep submicron regime. Higher temperature and temperature gradient also degrade the circuit performance as well as the reliability of the die and the package. Early considerations at different design levels of these thermal effects can lead to more accurate design parameter estimations and faster design convergence [2][3]. The ability to accurately model full-chip and package temperature distribution is thus required. A typical ceramic ball-grid array (CBGA) package consisting of the chip (die), thermal interface material, heat spreader, heat sink and other layers of packaging components is shown in Fig. 1. Through this package, heat can be removed from die to ambient air by two paths—a primary path from the die to heat spreader, heat sink and ambient; a secondary path from the die to the C4 pads and ceramic substrate and printed-circuit board. The primary path accounts for most of the heat removal from the die [4].

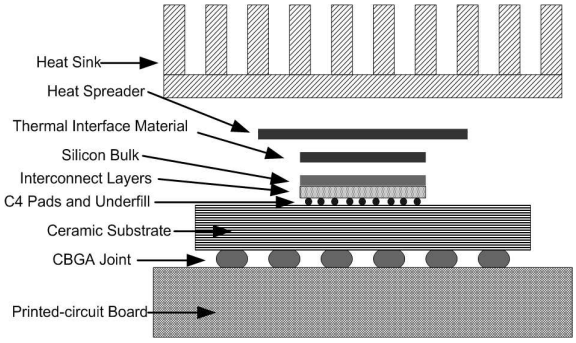


Figure 1: Packaging components in a typical CBGA package.

In recent years, there have been a number of existing die-level full-chip thermal models that provide detailed die temperature distributions. In [5][6][7], the authors present different detailed full-chip thermal models. These models all have detailed temperature distribution information across the silicon die and can be solved efficiently. Unfortunately, a limitation of the above models is that the thermal package is over-simplified. For example, the thermal interface material (TIM, the thermal paste between die and heat spreader) and heat spreader that greatly affect die temperature distribution are not included in the models. The bottom surface of the silicon substrate is also treated as isothermal, which significantly deviates from the real situation and therefore introduces errors. These simplifications of the thermal model are mainly due to the fact that these thermal models are developed by researchers in the circuit and die-level design areas.

In comparison, there are also several package-level thermal models [8][9][10]. These thermal models consist of simple networks of thermal resistances, whose values are extracted by data-fitting from the results of accurate but time-consuming detailed numerical package thermal model simulations (e.g. finite element method). Therefore, they are not fully parameterized and cannot be easily used to explore new package designs. In addition, these thermal models can provide only one or a few die-level temperatures, which is far from enough for fine-grained die-level designs.

A third category of existing thermal models includes the ones that model both die-level and package-level temperature distributions. Modeling detailed package temperature distribution is an important attribute for a thermal model to be more useful. This is because IC package components, especially the thermal interface material, heat spreader and heat sink, can greatly affect the die temperature and temperature distribution, as can be seen in Sec-

ISLPED'05, August 8–10, 2005, San Diego, California, USA

tion 3. Without modeling these components, a full-chip thermal model could lead to inaccurate temperature estimations, hence incorrect design decisions. In our previous work [3], a compact thermal model based on a grid-like lumped thermal R-C circuit is presented. This model can provide localized die-level full-chip temperature details as well as ways to model the temperature distributions of different packaging components for both the primary and secondary heat transfer paths. This thermal model can also be further extended to be flexible enough to model emerging packaging schemes such as stacked chip-scale packaging (SCP) [11] and 3D IC [12]. Throughout this paper, we adopt the thermal model presented in [3] to show that a thermal model able to model both chip and package temperatures is crucial to obtain accurate die temperature distributions.

The contributions of this paper are: (1) We explicitly demonstrate the significant impact of detailed temperature distribution on the accuracy of leakage power estimations by an example of thermally self-consistent leakage calculations for a POWER4-like microprocessor design. (2) We, for the first time, show the circuit design and computer architecture communities that modeling package details is an indispensable part for a die-level thermal model to be really useful. With several example thermal analyses regarding different properties of the thermal interface material (TIM), we show that omitting a package component in the thermal model can lead to significant errors in die temperature estimations. (3) Naturally, a full-chip and package thermal model can also act as a convenient medium for enhanced collaborations among circuit, architecture and package designers. This implies a design flow leading to early design evaluations from a thermal point of view. If potential thermal hazards are discovered early in the design process, different design tradeoffs can be carried out in an efficient way.

This paper is organized as follows. Section 2 shows an example of using the accurate full-chip and package thermal model to calculate across-die leakage power in a thermally self-consistent manner for a POWER4-like microprocessor design. Section 3 presents several example analyses showing the importance of modeling package components in addition to the silicon die, in particular, the thermal interface material. Then in Section 4, we present a thermally optimal design flow, which uses the full-chip and package thermal model as an efficient communication channel among circuit designers, computer architects and package designers to achieve a thermally optimized design in the early design stages. Finally, Section 5 concludes the paper and points out future work.

2. THERMALLY SELF-CONSISTENT LEAKAGE POWER CALCULATION

We first show as an example that using a full-chip and package thermal model, such as the one in [3], one can achieve accurate thermally self-consistent leakage power calculations for a POWER4-like microprocessor design at 130nm technology node. The leakage power calculation flowchart is shown in Fig. 2. Although similar leakage calculation methods have been described in [6] and [13], the literature still lacks for explicit data showing the impact of leakage power on die temperature distribution. In this section, we provide these data by showing the detailed die temperature maps with and without considering leakage power. We also show how the accuracy of temperature estimations impacts the accuracy of leakage power calculations by using temperature readings from the thermal model versus using a constant heuristic temperature across the die.

In this example, the floorplan of a POWER4-like microprocessor is generated by observing the real POWER4 floorplan. It is similar to the one shown in [14], except that the two cores are mirrored to each other, not simply duplicated. The full-chip and package compact thermal model is then constructed based on this floorplan and preliminary package data, which in a real design are from the package designers. For this example design, we use a packaging structure similar to the one in Fig. 1. In Fig. 3, we show the thermal model structure used for the primary heat transfer path (from silicon to heat spreader and to heat sink) of Fig. 1. The secondary heat

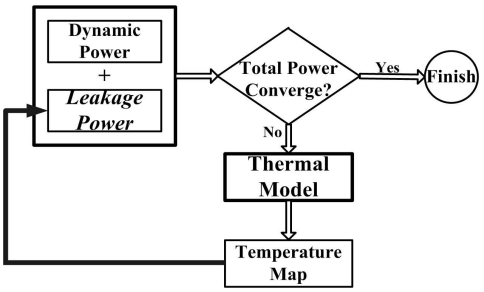


Figure 2: A full-chip thermal model closes the loop for accurate leakage power calculations. The loop is iterated until either power/temperature convergence is achieved or thermal runaway is detected.

transfer path from the die to C4 pads and to PCB is neglected since only a small amount of heat (less than 10%) is transferred through this path and including this path does not significantly change the results. Including the secondary path can be done by adding more layers of materials to the existing model, which will be part of our future work. In the thermal model, the silicon die, thermal interface material and the center part of the heat spreader that is covered by the thermal interface material are all divided into 40×40 grid cells to achieve detailed temperature distributions of these layers. (For clarity of illustration, only 3×3 grid cells are drawn in Fig. 3, although we really use 40×40 grid cells.) In order to validate the above-mentioned thermal modeling approach, besides the the validation work shown in [3], we have further quantitatively validated a similar thermal model for a real industrial design with a detailed ANSYS finite-element model simulations and qualitatively validated with on-chip temperature sensor measurements and infra-red temperature images for the same industrial design.

In order to get reasonably accurate initial power estimations of each functional unit for this POWER4-like microprocessor design, we combine IBM's cycle-accurate Turandot performance simulator [15] and PowerTimer power modeling tool [16] running benchmark program *bzip2*. The initial power inputs to the thermal model are the dynamic power values of running *bzip2* for each functional unit. The power numbers are further area-weighted into equivalent heat sources to each of the 40×40 grid cells. Leakage power of each functional unit is initially set to zero. After the thermal model is solved for the first time, leakage power of each unit is updated according to the updated temperature of that unit, which in turn updates the total power of the unit and changes the inputs to the thermal model and thus forms a loop as shown in Fig. 2.

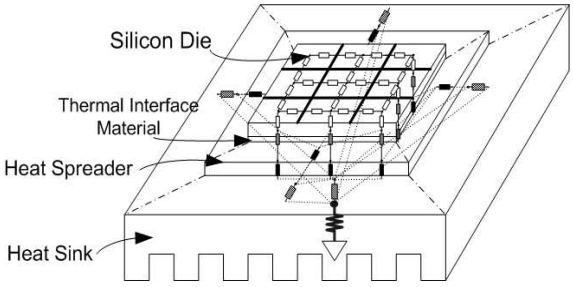
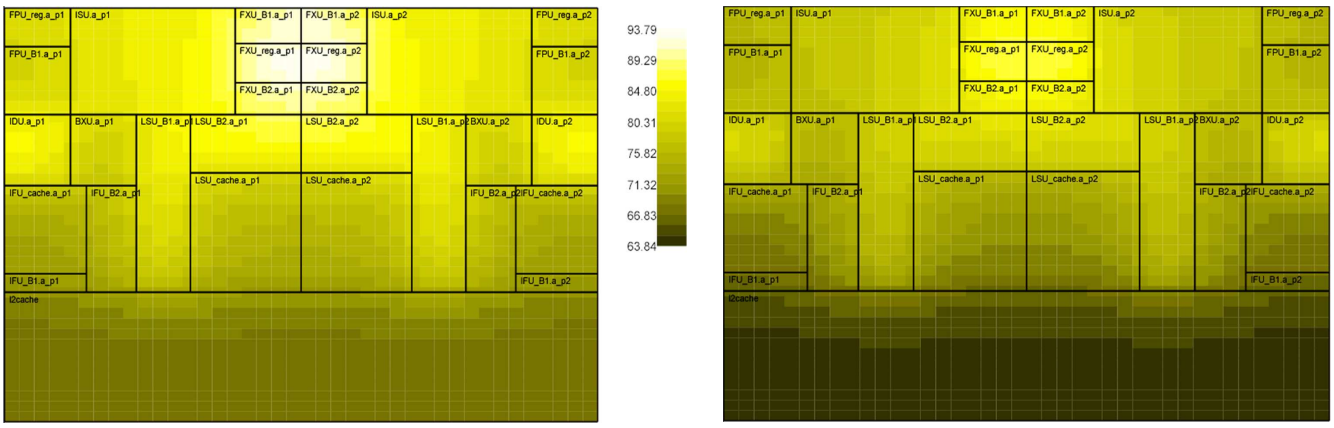


Figure 3: Example compact thermal model with 3x3 grid cells for silicon die [3]. For clarity of drawing, the structure is upside-down compared to the primary heat transfer path in Fig.1. Thermal capacitors and heat sources are also omitted for clarity.

The leakage power of grid cell i can be expressed as

$$P_{\text{leakage}_i} = A_i * \alpha * e^{\beta*(T_i - T_{base})}$$



(a) (b)

Figure 4: (a) Temperature map with considerations of the thermally self-consistent calculated leakage power for a benchmark workload on the POWER4-like microprocessor design at 130nm technology node. (b) Temperature map for the same design considering only dynamic power. Especially take note of the two FXU register files at the center of the top part, which are 7 degrees hotter in (a) compared to (b). (All temperatures are in Celsius.)

where A_i and T_i are the area and temperature of the grid cell. α and β are empirical factors that have different values for different technologies (e.g. $\alpha = 1{\times}10^5 \text{W/m}^2$ and $\beta = 0.025$ for 130nm). Typical values of α can be found in [17], and typical values of β can be found in [18]. T_{base} is the reference temperature at which α and β are defined.

The loop in Fig. 2 is iterated until the operating temperature and the total power converge, or thermal runaway is detected. For this design, convergence is usually achieved within 5 to 7 iterations with zero initial leakage power. On the other hand, if the leakage power is initialized according to room temperature, about 4 iterations are enough to achieve convergence. The computation time for solving a thermal circuit consisting 40×40 nodes is less than one minute on an AMD MP 1.5GHz system during each iteration. The computation time is proportional to the number of grid cells. If needed, further node reduction techniques, such as algebraic multigrid (AMG) method in [6], can be easily adapted to improve the computation time.

The major advantage of using the method in Fig. 2 is that it offers much more accurate leakage power calculations. Fig. 4(a) shows the converged temperature map which includes the effects from temperature-dependent subthreshold leakage power as well as dynamic power. For this particular design at 130nm technology node, the leakage power is 17.44% of the total chip power, and the temperature difference across the chip is 23.2°C for benchmark *bzip2*. Note that the locations of hot spots can change with different workloads. In comparison, Fig. 4(b) shows the temperature map with only dynamic power. Leakage power itself can raise the die temperature by 4 to 7 degrees for this 130nm design.

In addition, in order to decouple the temperature rise caused by the leakage power to that caused by the dynamic power, Fig. 5(a) shows the temperature map where only the thermally self-consistent leakage power is applied to the design. For comparison, Fig. 5(b) shows the temperature map where only non-temperature-aware leakage power is applied to the design, i.e., in Fig. 5(b), leakage power is calculated by assuming constant 85°C across the die. This is still the most common method for leakage power estimation in industry, but it is not accurate. As listed in Table 1, using heuristic constant temperature for leakage calculation for this design underestimates leakage power at the hottest spot by about 15.3%, and overestimates leakage power at the coolest spot by about 51.6%. In total, using heuristic constant temperature overestimates the overall chip leakage power by about 15.7%. (Notice that the colored temperature scale in Fig. 5 is not the same as the one in Fig. 4.) From these results, it is obvious that considering chip temperature variation with the actual temperature map from a thermal model is required for accurate leakage power calculation for this particular 130nm design, not to mention designs at future sub-100nm technologies. The inaccuracy of using constant heuristic temperature for the across-die leakage power calculations can cause unnecessary packaging cost, hence adding packaging cost (in the case of leakage overestimation); or put the chip in the danger of thermal hazards, hence lowering the final yield (in case of leakage underestimation).

scenarios	leakage and overall powers	
	hottest spot	coolest spot
actual thermal map	(FXU Regfile) leakage:0.144W overall:1.654W	(L2 Cache) leakage:2.406W overall:2.661W
85°C constant temp.	leakage:0.122W overall:1.632W	leakage:3.648W overall:4.263W

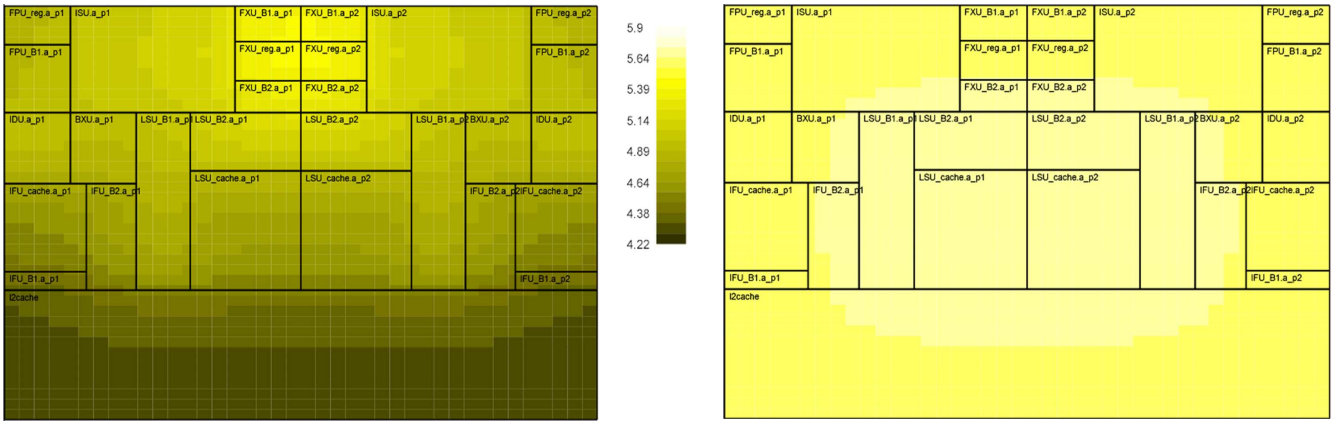
Table 1: Comparison of the leakage power values calculated using the actual temperature map from the thermal model to that calculated with a heuristic constant 85°C across the die, for both the hottest spot (FXU Regfile) and coolest spot (L2 Cache) on the die. (FXU Regfile is hotter because it has higher *power density* than L2 Cache, although its overall power is less than that of the L2 Cache.)

Using the thermally self-consistent leakage calculation method in Fig. 2 can also detect whether the design could possibly run into thermal runaway. One example is that if the preliminary package design doesn't have enough capability to dissipate all the generated heat, the loop in Fig. 2 can turn into a positive feedback loop. In this case, the leakage power and the temperature don't converge, hence thermal runaway occurs. The criterion for the occurrence of thermal runaway is indicated in [13][19] as

$$\frac{\partial^2 T}{\partial t^2} > 0$$

where T is temperature and t is time.

Table 2 shows the results of our investigation of potential thermal runaway for the 130nm process technology used in the example design. As can be seen, thermal runaway can be caused by elevated power dissipation of the design (from 55.74W to 139.35W) with the same thermal package as the example design. It can also be caused by defects in the package, e.g. voids or air bubbles in the thermal interface material, or imperfect attaching surface of the thermal interface material. Defects in the package can equivalently increase



(a) (b)

Figure 5: Imaginary temperature maps with only leakage power applied to the silicon, in order to isolate the thermal effect of leakage power. (a) Temperature map with the thermally self-consistent calculated leakage power applied for a benchmark workload on the POWER4-like microprocessor design. (b) Temperature map with leakage power calculated at constant 85°C across the silicon die. Notice the colored temperature scale in this figure is different from the one in Fig. 4. (Temperatures are in Celsius.)

the thermal resistance from the die to the ambient. As shown in Table 2, an increase in the equivalent package thermal resistance from 0.25°C/W to 0.8°C/W could result in thermal runaway for the design.

	No thermal runaway	thermal runaway
total power	55.74W	139.35W
package R_{th}	0.25°C/W	0.8°C/W

Table 2: For the 130nm process of the example design, an increases in total power from 55.74W to 139.35W, or an increase in equivalent junction-to-ambient thermal resistance from 0.25°C/W to 0.8°C/W can make thermal runaway happen.

As CMOS processes continue to scale into the sub-100nm regime, both operating temperature and leakage power increase significantly. To make things worse, industry usually tries aggressive techniques, such as controlling the gate length of a transistor to the lower-end of the gate-length variation (e.g. -1 to -3σ) in order to gain more transistor performance. This in turn exacerbates the leakage power consumption and makes thermal runaway of a design much easier to happen. Therefore, the accurate thermally self-consistent leakage power calculation method shown in Fig. 2 becomes more and more important for future technologies.

3. MODELING PACKAGE IN THE COMPACT THERMAL MODEL

In Section 1, we have mentioned that package components can greatly affect the temperature distribution across the silicon die. In this section, we show some example thermal analyses regarding one packaging component—thermal interface material (TIM) that glues the silicon die to the heat spreader. Other package components can be analyzed similarly.

With the flexibility of the compact chip and package thermal model in [3], we can easily investigate the thermal impacts of different TIM properties, such as its thickness, void size, and attaching surface roughness, in early design stages and provide important insights for circuit designers, computer architects and package designers.

We first show how the thickness of TIM affects silicon die temperature distribution. Fig. 6 plots the across-die temperature difference from the compact thermal model with different TIM thickness.

As can be observed from Fig. 6, thicker TIM results in poor heat spreading which leads to large temperature differences across the die. We can see that thick TIM can lead to very large die temperature difference across the die (>50°C). Even with nominal TIM thickness, which is 20μm for this design, the temperature difference across die is still 24°C. This means that the bottom surface of the die can not be modeled as an isothermal surface. If the TIM is thick enough, the resultant extremely large temperature differences across the die may be disastrous to circuit performance and die/package reliability. Using a better heat sink will only lower the average silicon temperature but will not help to reduce the temperature difference. This analysis suggests that using the thinnest possible TIM is one of the key issues for package designers to consider. On the other hand, with the known TIM thickness that can be best assembled in package with state-of-the-art packaging technology, it is the task of circuit designers and computer architects to design proper circuits and architectures to maintain the temperature difference across die within a manageable level.

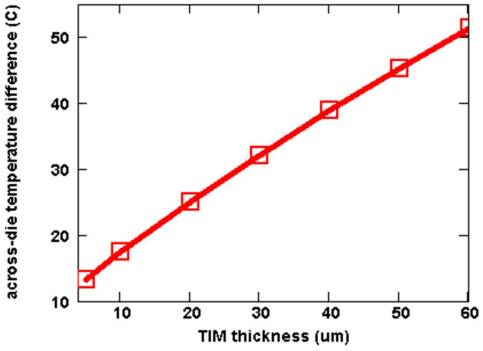


Figure 6: The impact of thermal interface material (TIM) thickness to silicon die temperature difference.

As another example, Fig. 7 shows the relationship between the size of TIM void and the hot spot temperature. During the packaging process, it is almost unavoidable to leave voids or air bubbles in the thermal interface material. In the compact thermal model, the void in TIM can be easily modeled by introducing higher vertical TIM thermal resistance to the grid cell where the void resides. Different sizes of the TIM void can be modeled by different sizes of the grid cell. For the simulations of Fig. 7, we put the TIM void right under the hottest grid cell, thus modeling the highest possible die temperature in the presence of a void with different sizes. As can be seen from Fig. 7, if the hot spot temperature of the design is

95°C, a void or air bubble in the TIM with a size of 0.25mm^2 can make the hot spot temperature drastically higher (290°C), which inevitably leads to thermal runaway of the chip. Therefore, it is desirable to improve the packaging techniques to make the size of the TIM void as small as possible. Package designers usually have the expertise to know typical TIM void sizes for different packaging processes. They can include this information in the thermal model. By doing this, the thermal model is now able to provide possible worst-case temperature regarding TIM void defects. The consequent architecture and circuit design decisions can thus avoid potential thermal hazards caused by the TIM void defects.

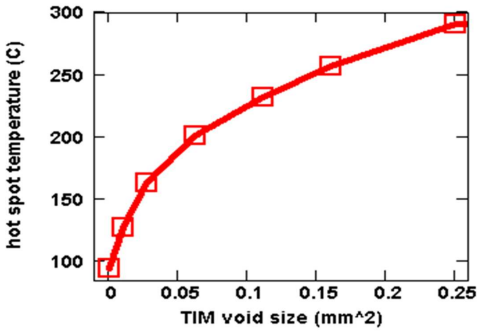


Figure 7: The impact of the size of void defect in thermal interface material (TIM) to silicon die hottest temperature. Temperatures are normalized to the ideal case where there is no void defect in the TIM layer. TIM void sizes are with the unit of mm^2.

Another important thermal interface material property that affects the die temperature is the surface roughness, i.e. non-uniform TIM. In real-life chip packaging process, the bottom surface of the die and the TIM's attaching surface cannot be perfectly smooth. As shown in Fig 8, TIM is only attached to the die at the bumps of the TIM surface. This causes ineffective heat conduction and hence higher die temperature comparing to the case where TIM and the die attach to each other perfectly. In order to investigate the impact of TIM non-uniformity to the die temperature, we change the thermal model of the TIM layer according to Fig 8, where we simply model the non-uniformity of the TIM surface as tiny bumps with spacing $2L$. The size of each grid cell is set to L. Therefore, heat can only be conducted through the grid cells representing the touching bumps. Grid cells representing the valleys are essentially tiny voids that do not touch the die and have extremely low thermal conductivity. The value of L thus can be used as an indicator of the non-uniformity of the TIM surface—the surface is rougher when L is larger and vice versa. Fig. 9 is the model results showing the relationship between L (non-uniformity) and die temperatures, where $L = 0$ means the TIM surface is perfectly uniform. As observed, even tiny non-uniform TIM surface (e.g. L=5μm) can significantly raise both the hottest and the average die temperature (by about 10 degrees). Package designers again usually have the specifications of the surface non-uniformities for different packaging processes. Without considering such package processing specifications, it is inevitable that a thermal model underestimates the die temperature and leads to designs that are not thermally optimized and designs with higher probability of premature failures.

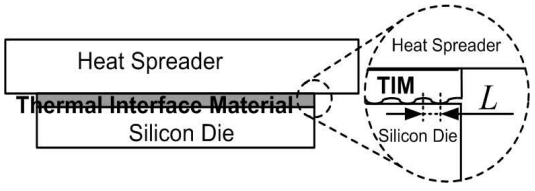


Figure 8: Close-up view of the TIM/die attaching surface. Surface non-uniformity is indicated by L.

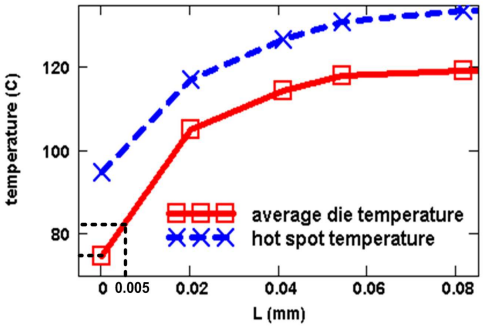


Figure 9: Hottest die temperature and average die temperature vs. the non-uniformity of TIM attaching surface. The larger L is, the rougher the attaching surface. L is defined in Fig. 8.

4. A THERMALLY OPTIMIZED DESIGN FLOW

From the above thermal analysis examples in Section 2 and 3, it is obvious that for optimal designs at future technologies, operating temperature needs to be modeled as accurate as possible in early design stages. In order to model temperature more accurately, important aspects of package information should also be included in the model. Ultimately, the full-chip and package thermal model should include all the needed package information (e.g. heat dissipation capability, geometries, materials, potential packaging defects such as the ones in Section 3, etc.) for different available package designs that circuit designers and architects can choose from and evaluate. Essentially, this requires more collaborations among circuit designers, computer architects and package designers. A compact thermal model that models detailed temperature distributions for both the silicon and the package can act as a convenient medium for such purpose. Fig. 10 illustrates a pre-layout design flow reflecting the collaborations among designers at different design levels. This design flow can detect potential thermal hazards early in the design process and lead to thermally optimized design.

As shown in Fig. 10, circuit designers first design basic circuit blocks called macros, and each macro has a simulated dynamic power for certain workload. It also has an estimated layout bounding box. The macros are then assembled into a preliminary microarchitecture and a floorplan according to work of computer architects. At this stage, initial total power, including rough estimation of leakage power, can be used for a package designer to propose a preliminary package design. All the information of power, floorplan and package are used to construct a compact thermal model which can perform thermally self-consistent leakage power calculations as shown in Section 2. The resulting temperature map can then be utilized to perform temperature-critical reliability analyses (e.g. interconnect electromigration, gate-oxide breakdown and package deformation) and temperature-related performance analyses (e.g. interconnect/device delay, power grid IR drop).

The results of all these analyses, together with the total powers, are then compared to the design goals. If the goals are not satisfied, different tradeoffs can be made—circuit designers may need to invent novel circuits with lower power dissipation, computer architects may think more about new architectures and different floorplans to better manage power and temperature, or package designers may need to propose more advanced, usually more expensive, packages. On the other hand, if the design goals are fully satisfied, we still need to check whether the design is too conservative and the design margin is too large for the application. We can then improve the conservative design by either introducing more aggressive circuit and/or architecture solutions to enhance performance, or using simpler and cheaper packages to reduce the cost of final product. These decisions and tradeoffs can then be evaluated

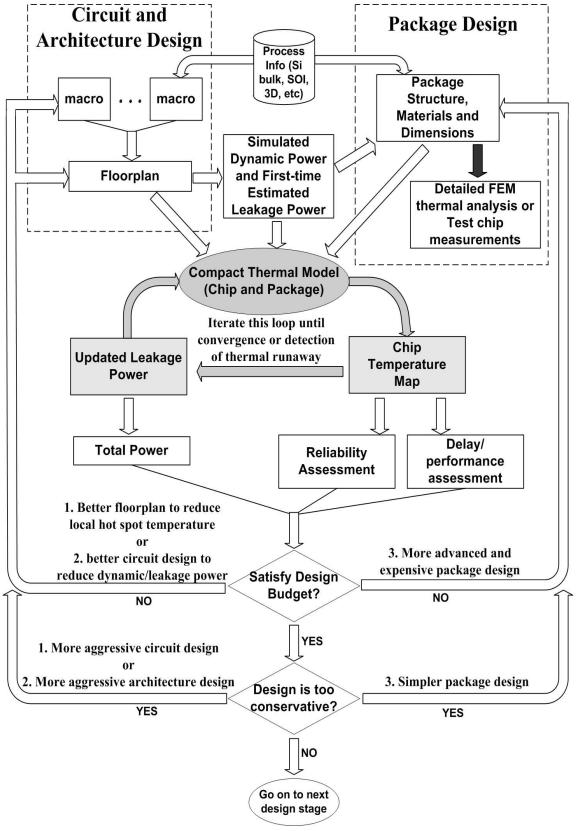


Figure 10: A design flow showing the compact thermal model acts as a convenient medium for productive collaborations for designers at the circuit, architecture and package levels.

by the thermal analysis again following the same flow until an optimal design point is reached. Then one can proceed to the physical design stage.

With the above design flow, the potential thermal hazards can be discovered and dealt with early and efficiently, thus the design is optimized from a thermal point of view.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we have shown example applications of an existing compact thermal model that provides detailed temperature distributions for both the silicon die and the package. From the thermally self-consistent leakage power calculations for an imaginary POWER4-like microprocessor design, we have seen that operating temperature needs to be modeled as accurately as possible. In order to achieve accurate temperature estimation, package components have to be also included in the thermal model, which has been demonstrated by several thermal analysis examples regarding thermal interface material in Section 3. Then it becomes clear that such a compact thermal model can act as a convenient communication channel for designers at the circuit, architecture and package levels. A temperature-aware design flow has also been proposed which is able to discover and solve potential thermal hazards in early design stages. With such a design flow, we can achieve a thermally optimized design.

There are other possible applications of this full-chip and package thermal model as future work. For example, regarding the impact of thermal package components to operating die temperature distribution, we can further investigate their impact on chip life time (transistor, interconnect and package) and extend existing reliability analyses [20][21]. We can also investigate the impact of random variations in package parameters (to name a few: TIM thickness, roughness, thermal conductivity, etc.) to the temperature, performance and power consumption of the chip design. These package parameter variations, together with the die-level parameter variations that are popularly considered (e.g. gate length, channel doping, oxide thickness, etc.), can greatly affect future sub-100nm circuit, architecture and package designs.

Acknowledgments

The authors would like to thank Scott Neely and Thomas Bucelot at IBM T. J. Watson Research Center for their invaluable discussions and suggestions, and Joshua Rosenbluh at Grinnell College, Iowa, for his information about generating SVG thermal plots. We would also like to thank the anonymous reviewers for their insightful comments. This work is supported in part by NSF grants CCR-0103364 (CAREER), CCF-0429765, Army Research Office grant W911NF-04-1-0288, a Univ. of Virginia FEST Award, two research grants from Intel MRL, and an IBM Faculty Partnership Award.

6. REFERENCES

[1] The international technology roadmap for semiconductors (ITRS), 2003.

[2] T. Kam, S. Rawat, D. Kirkpatrick, R. Roy, G. S. Spirakis, N. Sherwani, and C. Peterson. EDA challenges facing future microprocessor design. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 19(12):1498–1506, December 2000.

[3] W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan, S. Ghosh, and S.Velusamy. Compact thermal modeling for temperature-aware design. In *Proc. of 41st Design Automation Conference (DAC)*, pages 878–883, June 2004.

[4] J. Parry, H. Rosten, and G. B. Kromann. The development of component-level thermal compact models of a C4/CBGA interconnect technology: The motorola PowerPC 603 and PowerPC 604 RISC microproceesors. *Components, Packaging, and Manufacturing Technology–Part A, IEEE Transactions on*, 21(1):104–112, March 1998.

[5] T-Y. Wang and C. C-P. Chen. 3-D thermal-ADI: A linear-time chip level transient thermal simulator. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 21(12):1434–1445, December 2002.

[6] H. Su, F. Liu, A. Devgan, E. Acar, and S. Nassif. Full chip estimation considering power supply and temperature variations. In *Proc. of Intl. Symp. on Low Power Electronics and Design (ISLPED'03)*, pages 78–83, August 2003.

[7] P. Li, L. Pileggi, M. Asheghi, and R. Chandra. Efficient full-chip thermal modeling and analysis. In *Proc. of Intl. Conference on Computer-Aided Design (ICCAD-2004)*, November 2004.

[8] C. J. M. Lasance. Two benchmarks to facilitate the study of compact thermal modeling phenomena. *Components and Packaging Technologies, IEEE Transactions on*, 24(4):559–565, December 2001.

[9] M-N. Sabry. Compact thermal models for electronic systems. *Components and Packaging Technologies, IEEE Transactions on*, 26(1):179–185, March 2003.

[10] E. G. T. Bosch. Thermal compact models: An alternative approach. *Components and Packaging Technologies, IEEE Transactions on*, 26(1):173–178, March 2003.

[11] Intel's stacked chip scale packaging — http://www.intel.com/research/silicon/mobilepackaging.htm.

[12] K. Banerjee, S. J. Souri, P. Kapur, and K. C. Saraswat. 3-D ICs: A novel chip design for improving deep-submicrometer interconnect performance and systems-on-chip integration. *Proceedings of the IEEE*, 89(5):602–633, May 2001.

[13] L. He, W. Liao, and M. R. Stan. System level leakage reduction considering the interdependence of temperature and leakage. In *Proc. of 41st Design Automation Conference (DAC)*, pages 12–17, June 2004.

[14] Y. Li, D. Brooks, Z. Hu, and K. Skadron. Performance, energy, and thermal considerations for SMT and CMP architectures. In *Proc. of High Performance Computer Architecture (HPCA-2005)*, February 2005.

[15] M. Moudgill, J. D. Wellman, and J. H. Moreno. Environment for PowerPC microarchitecture exploration. *IEEE Micro*, 19(3):15–25, 1999.

[16] D. Brooks, P. Bose, V. Srinivasan, M. Gschwind, P. G. Emma, and M. G. Rosenfield. New methodology for early-stage microarchitecture-level power-performance analysis of microprocessors. *IBM Journal of Research and Development*, 47(5/6), 2003.

[17] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers. The impact of technology scaling on lifetime reliability. In *Proc. of The Intl. Conf. on Dependable Systems and Networks (DSN-04)*, June 2004.

[18] S. Heo, K. Barr, and K. Asanovic. Reducing power density through activity migration. In *Proc. of Intl. Symp. on Low Power Electronics and Design (ISLPED'03)*, pages 217–222, August 2003.

[19] W. Liao and L. He. Coupled power and thermal simulation and its application. In *3rd Workshop on Power-Aware Computer Systems*, December 2003.

[20] Z. Lu, W. Huang, J. Lach, M. Stan, and K. Skadron. Interconnect lifetime prediction under dynamic stress for reliability-aware design. In *Proc. of Intl. Conference on Computer-Aided Design (ICCAD-2004)*, November 2004.

[21] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers. The case for lifetime reliability-aware microprocessors. In *Proc. of The Intl. Symp. on Computer Architecture (ISCA'04)*, June 2004.