# Fast Evaluation Method for Transient Hot Spots in VLSI ICs in Packages

Je-Hyoung Park
*UC Santa Cruz*
*jhpark@soe.ucsc.edu*

Ali Shakouri
*UC Santa Cruz*
*ali@soe.ucsc.edu*

Sung-Mo Kang
*UC Merced*
*smk123@ucmerced.edu*

## Abstract

*Recently VLSI IC design is concerned with the large temperature non-uniformity in high power chips. Thus far, thermal simulations have been limited to steady-state worst case conditions, which have caused the use of conservative margins in thermal designs. Transient temperature characteristics were not simulated in prior art chip-level simulations due to the high computational expense. To drastically reduce the time for the chip-level thermal simulations, we have developed a matrix convolution technique, called the Power Blurring (PB) method. Our method renders the temperature profile of a packaged IC with maximum error less than 3% for several case studies done and reduces the computation time by a factor of 100, compared to the simulations done by the industry standard finite element tools.*

## 1. Introduction

As CMOS VLSI technology scales down, increasingly more devices are integrated in a smaller chip area. Figure 1 shows the International Technology Roadmap for Semiconductors (ITRS) [1] prediction of microprocessor power density for the next decade. This figure conveys two important messages. First, due to the aggressive scaling down of CMOS technology, the power density of high-performance microprocessor continues to increase. In year 2015, the maximum power can reach 200W. Secondly, the power density of cost-performance microprocessor saturates in year 2008. The maximum power is projected to be close to 151W (double check these numbers). If we fail to handle power dissipation in ICs appropriately, the scaling of CMOS VLSI technology could be slowed down significantly.

The increased power density leads to more stringent requirements for the heat sinks. Non-uniform activities in different functional units yield non-uniform surface temperature distribution in a chip.

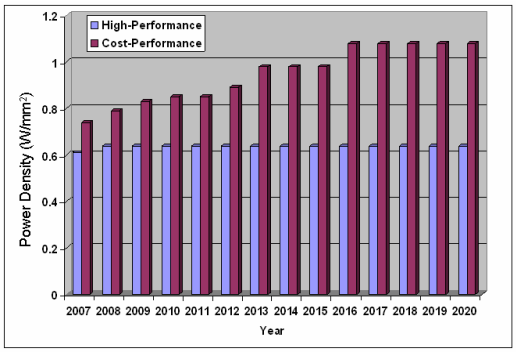



**Figure 1. The ITRS prediction for the power density of microprocessors**

Hot spots and non-uniformly elevated temperatures limit both the performance and the reliability of packaged IC chips. Temperature non-uniformity evolves with time and there exist transient hot spots. The transient temperature spike or localized heating can cause timing errors or even physical damages, leading to reliability failures. However, thermal simulations and design optimizations have been done under steady-state worst case conditions due to the high computational expense. This leads to the use of conservative margins in thermal designs. As the thermal budget becomes increasingly tighter, such approach has become too costly and ineffective. Fine granularity in the state-of-the-art chip-level transient thermal simulations is such that the inclusion of any realistic package configuration requires too many nodes and long simulations. This is computationally too expensive for physical design optimization or performance verification in a packaged environment.

In an effort to reduce the computation cost, the alternating-direction-implicit (ADI) method was proposed [2]. In ADI approach complicated package and heat sink structures are modeled by 1-D equivalent thermal resistance network. Ignoring the lateral heat spreading in realistic packages can result in overestimated chip temperatures [3].

To drastically reduce the time for chip-level steady-state and transient thermal simulations including a realistic package model, we developed a matrix convolution technique, called the Power Blurring (PB) method [4]. The remainder of this paper is organized as follows. In Section 2, the PB method will be discussed with steady-state case studies. Our transient simulation method and results for various case studies are presented in Section 3, followed by conclusions in Section 4.

## 2. Power Blurring Method

### 2.1. Full-chip package model

Figure 2 shows our full-chip package thermal model. The configuration consists of a Si IC with a surface area of 1cm× 1cm and a Cu heat sink with a heat spreading layer. In a flip chip package, most of the heat is considered to flow through the bottom surface of the heat sink, and other minor heat transfer paths are neglected.

### 2.2. The thermal mask

The temperature distribution of the IC chip can be regarded as a blurred image of the power map when it is blurred with an appropriate filter mask. In our case, the filter mask is the impulse response, which can be obtained by using a Finite Element Analysis (FEA) tool such as ANSYS [5]. Figure 3 shows the mask with unit of °C/W (thermal resistance).

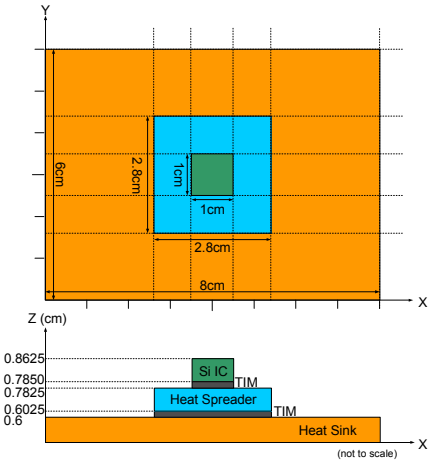



**Figure 2. Schematic of a packaged IC chip, where the heat spreader, the heat sink and the thermal interface material (TIM) are included.**

The thermal profile corresponding to a given power map is obtained when the thermal mask is convolved with the given power map. The advantage of the PB technique is that it can be applied to realistic chip geometries where analytical Green's function can not be obtained.

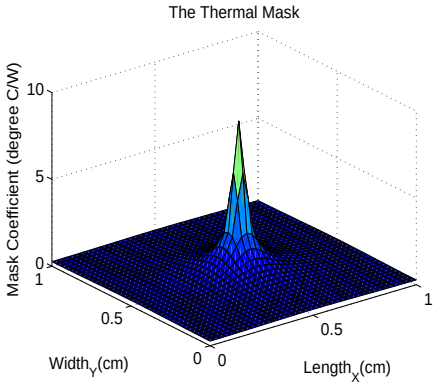


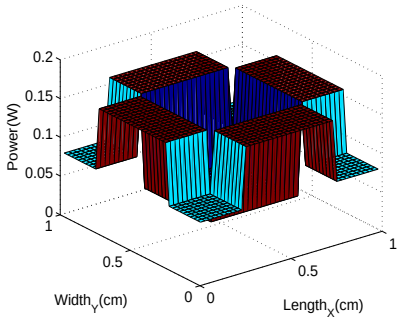

**Figure 3. The thermal mask**

### 2.3. Steady-state simulation results

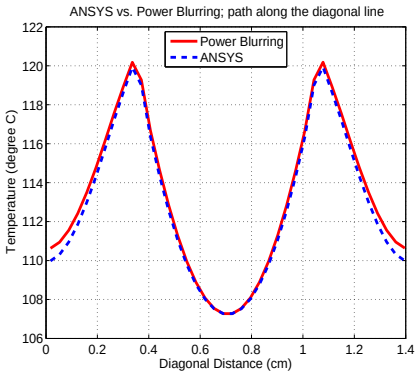
Figure 4 (a) shows the power map and Figure 4 (b) the corresponding temperature profile along the diagonal path. For comparison, both PB computation results and ANSYS simulation results are plotted together. Figure 4 (c) depicts the relative deviation with respect to ANSYS simulation result. The relative deviation is given by

$$E_r = \frac{|T_{PB} - T_{ANSYS}|}{T_{ANSYS}} \quad (1)$$

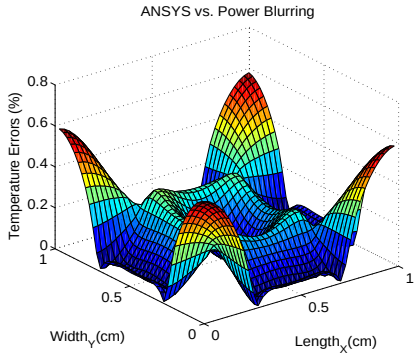
As can be seen, the PB method gives steady-state computation result that is in a good agreement with the ANSYS simulation result. The maximum temperature error is less than 1%.




(a)




(b)




(c)

**Figure 4. A case where high power consumption is at the side edges; (a) the power map (b) the temperature profile along the diagonal path, and (c) the relative deviation with respect to ANSYS simulation result.**

## 3. Transient thermal simulation

### 3.1. Simulation method

For transient simulations, time evolution of the thermal mask resulting from temporal and spatial delta function needs to be obtained.

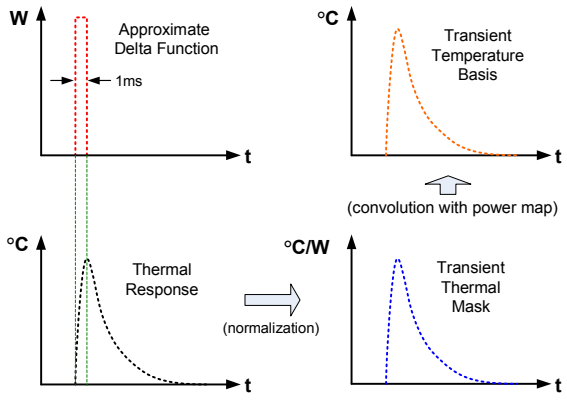



**Figure 5. Obtaining transient thermal mask and transient temperature basis (the visualization is given for a single spot on a chip.)**

In practice, a point heat source is applied for a short time period (e.g., 1ms) and the corresponding thermal (temperature) response is acquired at each time step (e.g., 0.2ms). The resulting thermal response is normalized with respect to the amount of applied power to create transient thermal mask. The transient thermal mask is convolved with a given power map at every time step to create a transient temperature basis as shown in Figure 5. To calculate a transient thermal profile for a long pulse input, the transient temperature basis is shifted in time domain for the pulse duration. The overall transient temperature is the aggregate of those transient temperature bases. The process is depicted in Figure 6.

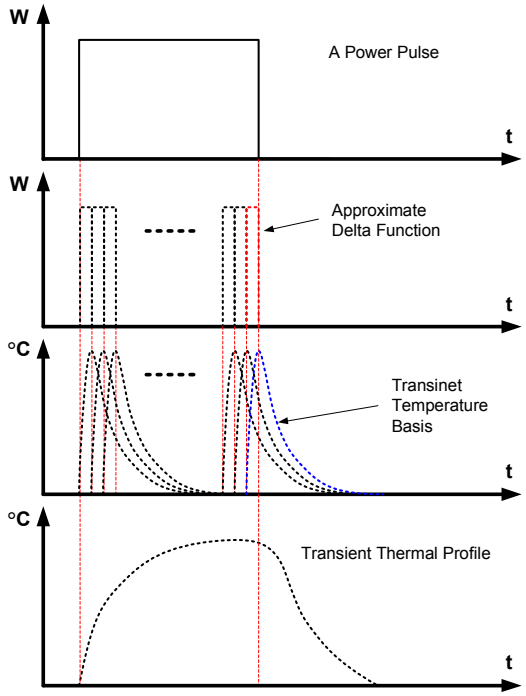



**Figure 6. Schematic overview of transient thermal simulation method; a long power pulse can be regarded as an aggregate of approximate delta functions. The overall transient thermal profile can be obtained by superposition of shifted transient temperature bases.**

### 3.2. Case studies

Two different coarse power maps ($P_1$ and $P_2$) in Figure 7 (a) are applied at t=0.1 sec and t=0.3 sec for 0.1 sec duration, respectively. Resulting temperature profiles at t=0.2 sec and 0.3002 sec are presented in Figure 7 (c) and (d), respectively. The transient temperature profile at the center of the IC chip is also presented in Figure 8.

$P_1$

| | | | |
|---|---|---|---|
| 4W | 4W | 4W | 4W |
| 4W | 16W | 16W | 4W |
| 4W | 4W | 8W | 6W |
| 4W | 4W | 4W | 4W |

$P_2$

| | | | |
|---|---|---|---|
| 8W | 4W | 2W | 4W |
| 4W | 6W | 8W | 4W |
| 4W | 4W | 16W | 9W |
| 6W | 4W | 4W | 4W |

(a)

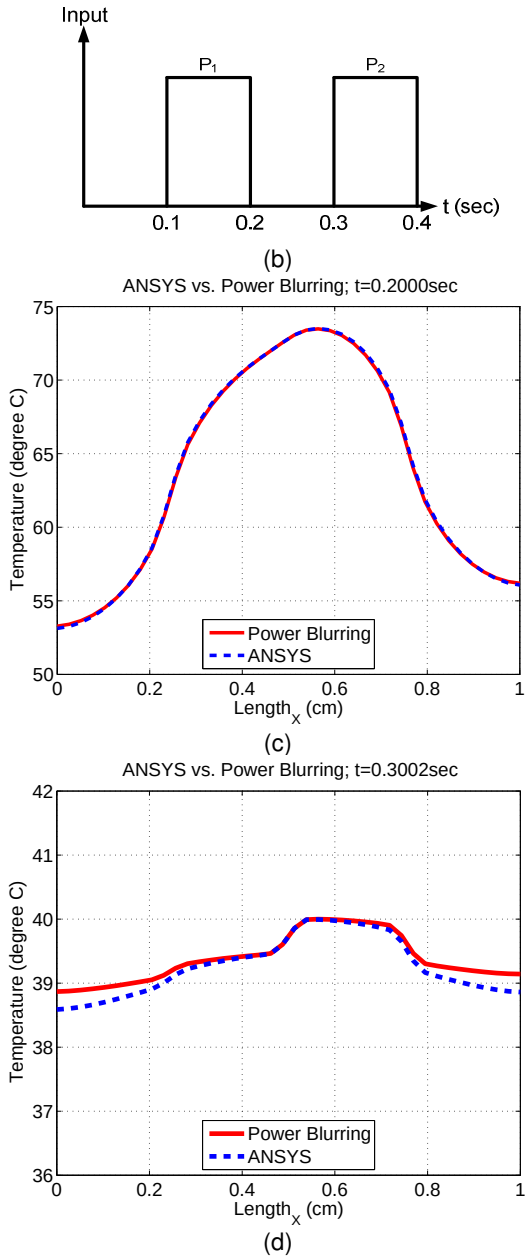



**Figure 7. Pulse input of the coarse power maps; (a) coarse power maps (b) power dissipation pattern (c) temperature profile along y=0.5cm at t=0.2 sec and (d) t=0.3002 sec**

For the last example, ANSYS simulation took 10,056 seconds whereas the PB method took only 100 seconds (reduction factor of 101). As can be seen, calculation results show good agreements with ANSYS simulation results. The maximum temperature error of 3% was observed in the transient case study.

## 4. Conclusion

In this paper, an accelerated temperature calculation method, Power Blurring (PB), was applied to various types of power maps. The PB technique yields the temperature profile in a packaged IC chip with maximum error less than 3% for all case studies done and reduces the computation time by a factor of 100, compared to the simulations done by the industry standard tool, ANSYS.

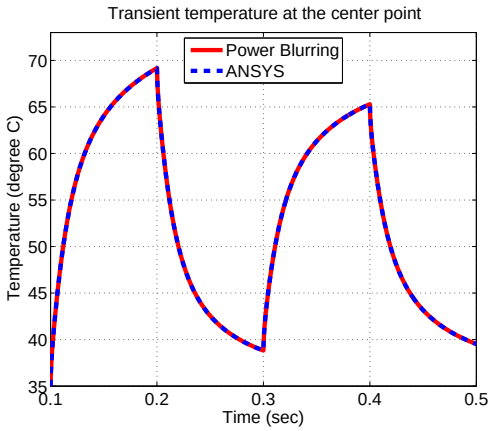



**Figure 8. Transient temperature profile at the center of the IC chip for the pulse input of the two different coarse power maps in Figure 7 (a)**

## 5. Acknowledgement

This work was supported in part by the National Science Foundation CSR SMA program.

## 6. References

[1] International Technology Roadmap for Semiconductors (ITRS), Semiconductor Industry Association, 2006 update.

[2] Ting-Yuan Wang and Charlie Chung-Ping Chen, "Thermal-ADI – A Linear-Time Chip-Level Dynamic Thermal-Simulation Algorithm Based on Alternating-Direction-Implicit (ADI) Method," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 11, No. 4, August 2003.

[3] Sheng-Chih Lin and Kaustav Banerjee, "An Electrothermally-Aware Full-Chip Substrate Temperature Gradient Evaluation Methodology for Leakage Dominant Technologies with Implication for Power Estimation and Hot-Spot Management," International Conference on Computer-Aided Design, San Jose, CA, November 5-9 2006.

[4] Travis Kemper, Yan Zhang, Zhixi Bian, and Ali Shakouri, "Ultrafast Temperature Profile Calculation in IC chips," Proceedings of 12th International Workshop on Thermal investigations of ICs (THERMINIC), Nice, France, pp. 133-137, 2006.

[5] ANSYS R11.0, Swanson ANSYS Inc., 2007