

## THERMAL IMPACTS OF LEAKAGE POWER IN 2D/3D FLOORPLANNING<sup>\*,†</sup>

YUCHUN MA<sup>‡</sup>, QIANG ZHOU, PINGQIANG ZHOU  
and XIANLONG HONG

*Tsinghua National Laboratory for Information Science and Technology (TNList),  
Department of Computer Science and Technology, Tsinghua University,  
Beijing, 100084, P. R. China  
<sup>‡</sup>myc@mail.tsinghua.edu.cn*

Received 26 September 2008

Accepted 23 April 2010

Leakage power is becoming a key design challenge in current and future CMOS designs. Due to technology scaling, the leakage power is rising so quickly that it largely elevates the die temperature. In this paper, we deeply investigate the impact of leakage power on thermal profile in both 2D and 3D floorplanning. Our results show that chip temperature can increase by about 11°C in 2D design and 68°C for 3D case with leakage power considered. Then we propose a thermal-driven floorplanning flow integrated with an iterative leakage-aware thermal analysis process to optimize chip temperature and save leakage power consumption. Experimental results show that for 2D design, the max chip temperature can be reduced by about 8°C and the proportion of leakage power to total power can be reduced from 19.17% to 11.12%. The corresponding results for 3D are 60°C temperature reduction and 16.3% less leakage power proportion.

*Keywords:* 3D ICs; thermal; leakage power; floorplanning.

### 1. Introduction

In CMOS digital circuits, power dissipation consists of dynamic and static components. In circuits with a high supply voltage, a relatively high transistor threshold voltage can be used, making subthreshold current negligible, which is the main component of leakage current. However, with continuous shrinking of minimal feature size, leakage power is expected to become a major challenge for future CMOS designs. Although leakage is about 10%–20% of total chip power for the current

\*Supported by the National Natural Science Foundation of China under Grant Nos. 60606007 and 60833004, Tsinghua Basic Research Fund JC20070021 and TNList Cross-discipline Foundation.

†This paper was recommended by Regional Editor Krishna Shenai.

‡Corresponding author.

generation of CMOS technologies, the number is expected to rise to 50% for the next generation technologies.<sup>1,2</sup>

As indicated in Ref. 3, the sub-threshold current of a transistor has a super-linear dependency on temperature. With the development of VLSI technologies, the size of a transistor is even smaller and hence the sensitivity of leakage power due to temperature is even more evident. Besides, the temperature of a block in a CMOS circuit is not confined to the block itself and it affects the temperatures of all its neighboring blocks due to thermal diffusion. Thus the temperature of a certain block cannot be determined without considering the floorplan of all the blocks. Gupta *et al.*<sup>4</sup> showed that different floorplans have different thermal profiles, thus have different leakage power. However, though previous floorplanning algorithms<sup>7-18</sup> have already taken thermal effects into consideration, few of them considered the effects caused by leakage power. Omitting the thermal-dependent leakage power in thermal analysis can result in significant temperature estimation errors. Therefore, it is very necessary to include the leakage power calculation in early design stage to get the accurate temperature profile.

The above observations motivate us to investigate the impact of leakage power on thermal profile in floorplanning, especially 3D floorplanning, in which thermal effect is one of the most important issues. In this paper, we first study the impact of leakage power on thermal profile in 2D and 3D floorplanning, and then propose a thermal-driven optimization flow to optimize the chip temperature and leakage power dissipation. The rest of the paper is organized as follows. In Sec. 2, we review the related work in leakage power modeling and thermal-aware floorplanning/placement. In Sec. 3, we present the thermal and leakage power models. Section 4 formulates our thermal-driven floorplanning flow with temperature-dependent leakage power considered. Section 5 shows the experimental results and the conclusions are provided in Sec. 6.

## 2. Related Work

The related work can be divided into two classifications: temperature-dependent leakage power modeling and thermal-aware floorplanning/placement.

Chandrakasan *et al.*<sup>3</sup> showed that the subthreshold current for a transistor is a super exponential function of its supply voltage, threshold voltage and temperature. Nassif *et al.*<sup>1</sup> proposed a model to estimate the leakage power for a chip while considering power supply and temperature variations. Liao *et al.*<sup>5</sup> also proposed a temperature-dependent leakage power model with an assumption that the whole chip is at one uniform temperature. Huang *et al.*<sup>6</sup> used a simple exponential model to express the relation between leakage power and temperature without loss in accuracy.

In the area of 2D thermal-aware floorplanning, Sankaranarayanan *et al.* have proposed thermal-aware floorplanning at the micro-architectural level for processors.<sup>7,8</sup> Han *et al.*<sup>9</sup> conducted a study on the floorplanner's effectiveness in lowering

the maximum processor temperatures. Recently, Healy *et al.*<sup>10</sup> have proposed thermal-driven and communication profile-driven floorplanning to reduce peak temperatures. Hung *et al.*<sup>11</sup> used floorplanning with genetic algorithms to reduce the peak temperatures. Reduction of clock tree power using activity based register clustering and thermal-aware placement has been proposed by Cheon *et al.*<sup>12</sup> and Obermeier *et al.*,<sup>13</sup> respectively.

Besides, Ekpanyapong *et al.*<sup>14</sup> and Cong *et al.*<sup>15</sup> have proposed thermal-driven and communication profile-driven 3D floorplanning. Hung *et al.* proposed thermal-aware floorplanning for 3D microprocessors,<sup>16</sup> the power consumption of interconnect is considered during floorplanning. Recently, Li *et al.* developed a hierarchical 3D floorplanning algorithm.<sup>17</sup> Goplen and Sapatnekar<sup>18</sup> developed force-directed thermal-aware standard-cell placement algorithms. These techniques demonstrate good and scalable performance. But none of them consider the influence of leakage power which may cause significant error. Therefore, it is very necessary to investigate how much the leakage power will influence the chip temperature on both 2D and 3D designs.

### 3. Background

This section will describe the thermal model and the temperature-dependent leakage power model.

#### 3.1. Compact thermal model

Thermal analysis is the simulation of heat transfer through heterogeneous material among heat producers (e.g., transistors) and heat consumers (e.g., heatsinks attached to an IC). Modeling thermal conduction is analogous to modeling electrical conduction, with thermal conductivity corresponding to electrical conductivity, power dissipation corresponding to electrical current, and temperature corresponding to voltage.<sup>19</sup> The full-chip and package compact thermal model is constructed based the floorplan and preliminary package data. As shown in Fig. 1, the circuit stack is first divided into tiles. A tile stack is modeled as a resistive network. The isothermal bases of room temperature are modeled as a voltage source. A current source is present at every node in the network to represent the heat sources. The tile stacks are connected by lateral resistances.

One can spatially discretize the system and solve the following equation to determine the steady-state thermal profile as a function of power profile.

$$T = PA^{-1}, \quad (1)$$

where  $A$  is an  $N \times N$  sparse thermal conductivity matrix.  $T$  and  $P(t)$  are  $N \times 1$  temperature and power vectors. Li *et al.* proposed a multi-grid thermal analysis technique that automatically refines spatial discretization to achieve orders of magnitude speedup compared to techniques using homogeneous partitioning without

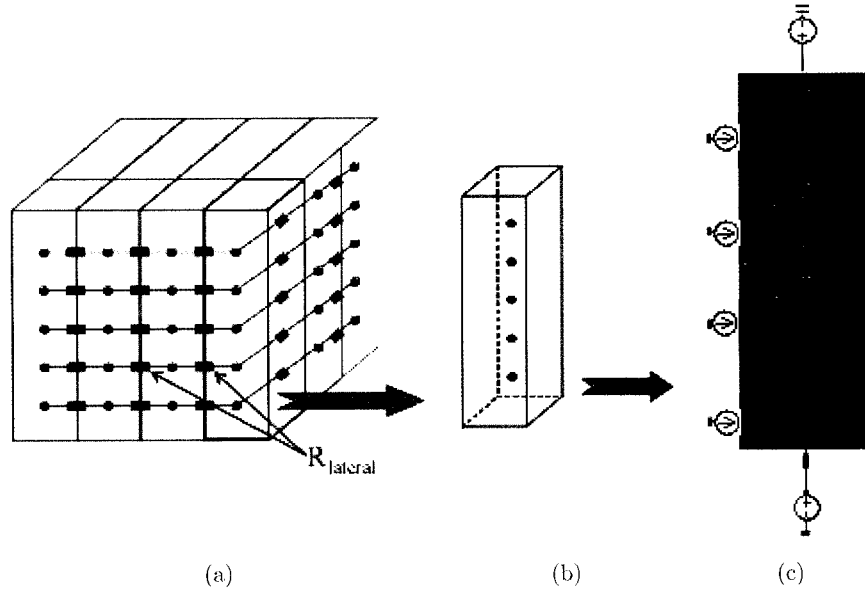


Fig. 1. Resistive thermal model for a 3D IC: (a) tiles stack array; (b) single tile stack; (c) tile stack analysis.

loss of accuracy.<sup>20</sup> We use this thermal solver and its extended multi-layer version. This thermal model uses incremental refinement to generate a tree of heterogeneous parallelepipeds thereby supporting fast thermal analysis without loss in accuracy.

### 3.2. Temperature-dependent leakage power model

The leakage power of a grid cell  $i$  (we divide each tile into a certain number of grid cells in the  $Z$  direction) can be expressed as

$$P_{leakage_i} = A_i \times \alpha \times e^{\beta \times (T_i - T_{base})}, \quad (2)$$

where  $A_i$  and  $T_i$  are the area and temperature of the cell.  $\alpha$  and  $\beta$  are empirical factors that have different values for different technologies (typical values found in Ref. 6 are  $\alpha = 1 \times 10^5 \text{ W/m}^2$  and  $\beta = 0.025$  for 130 nm with  $V_{dd} = 1.1 \text{ V}$ ).  $T_{base}$  is the reference temperature at which  $\alpha$  and  $\beta$  are defined. Figure 2 shows the relation between temperature and leakage power. The leakage power increases with the temperature super-linearly.

## 4. Thermal-Driven Floorplanning with Leakage Power

It is obvious that for optimal designs at future technologies, on-chip temperature needs to be modeled as accurate as possible in early design stages. Hence, thermal-driven floorplanning should take temperature dependent leakage power effect into consideration. Especially in 3D designs, the temperature distribution varies from layer to layer. The floorplan has a direct effect on the leakage power and the temperature

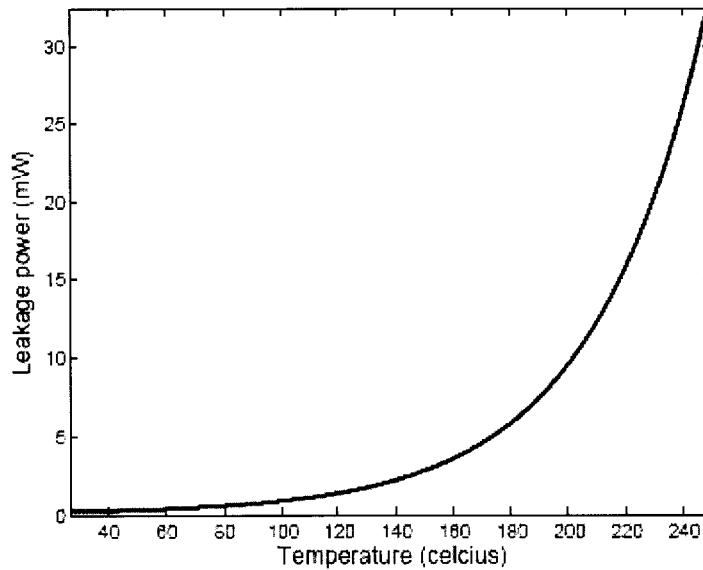


Fig. 2. Relation between leakage power and temperature.

dependent leakage power also influences the hotspots distribution on floorplans. To optimize the packing with leakage power considered, we need to design iterative thermal estimation and optimization flow to get the convergence.

**4.1. Leakage-aware iterative thermal analysis**

As shown in Sec. 3, leakage power and temperature values are interacting with each other. To get the accurate value of the temperature on chip caused by both dynamic and leakage power, we need the iterative computation process as shown in Fig. 3. The temperature profile is used to calculate the corresponding leakage power using leakage power model shown in Sec. 3.2. The total power distribution will be updated with the new leakage power. And the thermal distribution can be further updated

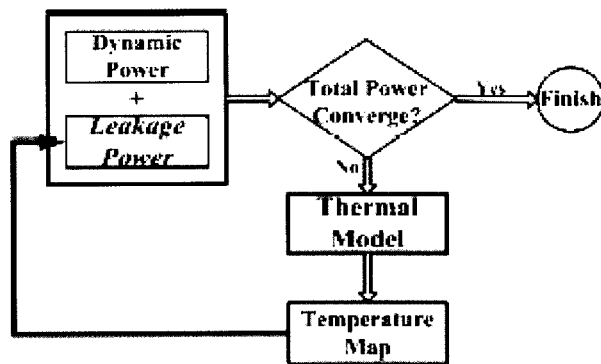


Fig. 3. Full-chip thermal model closes the loop for accurate leakage power calculation.

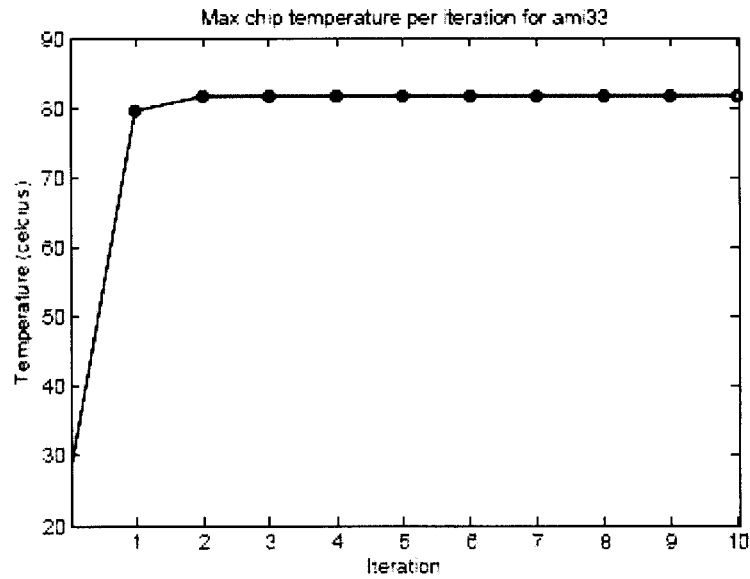


Fig. 4. Iteration process in thermal analysis.

using the thermal model in Sec. 3.1. The loop is iterated until either power/temperature convergence is achieved or thermal runaway is detected. It may need several iterations to achieve the convergence. As shown in Fig. 2, the temperature has small impact on leakage power when the temperature is low. To save the early iterations, the leakage power can be initialized according to room temperature. We show the iteration process for ami33 in Fig. 4. The iterations can be saved to 2 or 3 iterations to get an accurate estimation.

In order to show the temperature rise caused by the temperature-dependent leakage power, we first generate a 2D packing for a typical circuit ami33, as shown in Fig. 5. Figure 5(a) shows the corresponding dynamic power distribution. Then we use the thermal tool introduced in Sec. 3 to get the thermal profiles for this packing without and with leakage power consideration, as shown in Figs. 5(c) and 5(d). As shown in Fig. 5, the temperature with leakage power considered is much hotter than the temperature with only dynamic power. The thermal distributions for two cases are slightly different with each other since the hot regions in Fig. 5(d) are larger than those in Fig. 5(c). The temperature difference for a certain grid in these two pictures is about  $10^{\circ}\text{C}$  on average.

#### 4.2. Thermally optimized floorplanning

Different blocks in packings have different dynamic power dissipation profiles and hence produce varying local temperatures. The placement has a direct effect on the leakage power of the design. A leakage-aware floorplanner considering temperature profile is introduced in this section. The cost function of a leakage aware floorplanner not only includes the traditional objectives such as chip packing area and wirelength

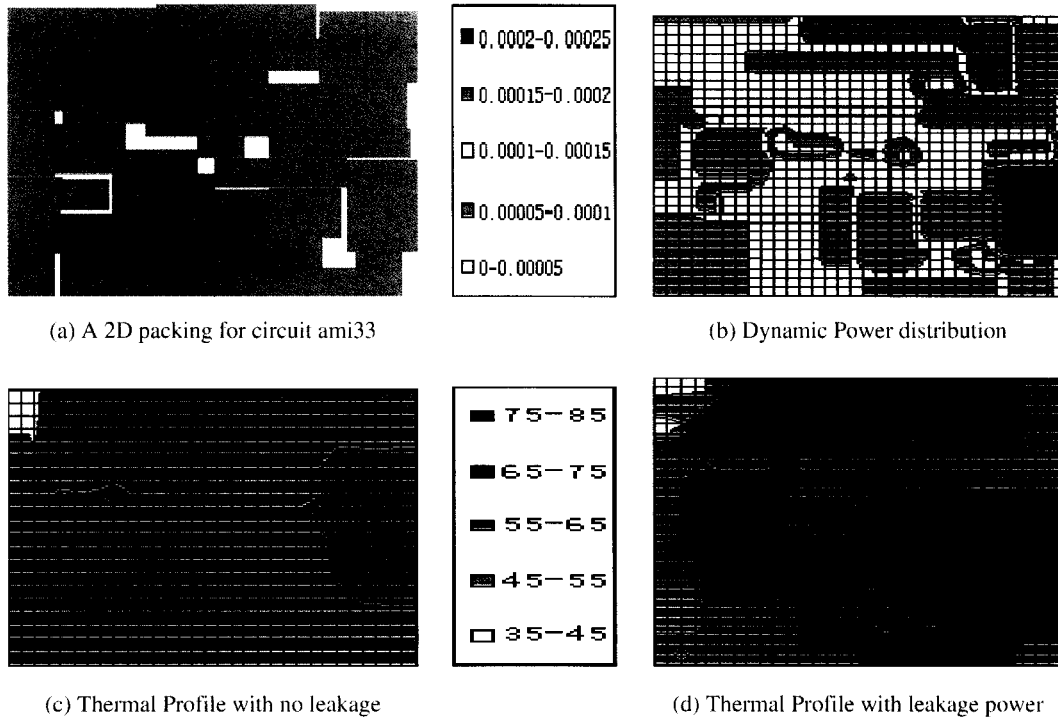


Fig. 5. Temperature maps of a 2D packing instance — ami33.

but also takes leakage power aware temperature into account. As shown in Fig. 6, thermally optimized floorplanning flow is an iterative improving process. Starting with an initial solution, we alternately do new solution generation and leakage aware thermal profile generation until no better solution can be found.

In this paper, we use a force-directed approach to implement the thermal-driven floorplanning. Any other traditional floorplanner can be extended to be a leakage-aware floorplanner with our model. The force-directed placement algorithm simulates the mechanics problem in which articles are attached to springs and their movements obey the Hooke's law. Here we use the basic force-directed engine derived from Ref. 22 and extend it to deal with 3D thermal driven floorplanning problem. We choose force-directed placement engine because of its good performance and scalable feature.

## 5. Experiment Results

Our algorithm is implemented in C++, and the experiments are performed on a workstation with 3.0 GHz CPU and 4 GB physical memory. We use five typical MCNC and GSRC benchmarks<sup>15</sup> in our experiments. We assign a power density of random number between  $1.0 \times 10^5$  (W/m<sup>2</sup>) and  $5.0 \times 10^6$  (W/m<sup>2</sup>) to each block in the above mentioned circuits, which are typical values in modern CMOS circuits. In 3D packing, *four* device layers are used for all circuits. The units in the following

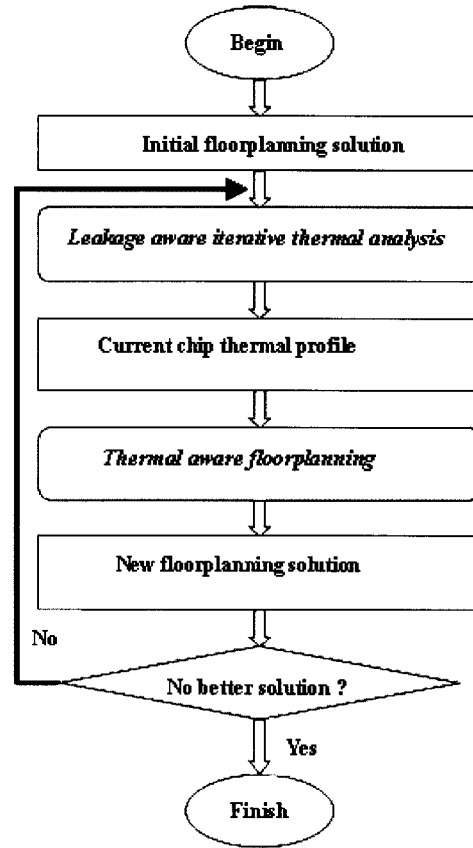


Fig. 6. Thermally optimized floorplanning flow.

results are Area in  $1.0 \times 10^4 \mu\text{m}^2$ , Wirelength (HPWL) in mm, Temperature in  $^{\circ}\text{C}$ , and Time in seconds (s).

### 5.1. Thermal impact of leakage power

Tables 1 and 2, respectively, show the impact of leakage power on thermal profile in 2D and 3D floorplanning. Here, we compare the results of temperature with fixed leakage power (FLP, which equals to the value at  $27^{\circ}\text{C}$  obtained by the model described in Sec. 3.2) and temperature-dependent leakage power (TDLP). As stated before, each 3D floorplan has *four* active layers.

From Table 1, we can see that for the same packings, the average and max chip temperatures, respectively, increase by about  $6.6$  and  $11.2^{\circ}\text{C}$  when we take temperature-dependent leakage power into account. In 3D floorplanning, however, the max temperature increases averagely by  $67.7^{\circ}\text{C}$ , which is about five times worse than the number in 2D case. The leakage power and temperature value are interacting with each other and the leakage power increases with the temperature super-linearly. Since the temperature on 3D chip are more than  $100^{\circ}\text{C}$ , and the temperature



Table 1. Thermal impact of leakage power in 2D case.

Circuit	Area	HPWL	Temperature (with FLP)		Temperature (with TDLP)		Time
			Avg	Max	Avg	Max	
Ami33	123.2	58.3	45.0	70.2	49.7 (+ 4.7)	81.6 (+ 11.4)	4
Ami49	3663.0	861.6	38.6	66.8	41.7 (+ 3.1)	77.8 (+ 11.0)	10
N100	18.8	191.0	52.8	69.3	60.0 (+ 7.2)	80.0 (+ 10.7)	24
N200	19.4	367.1	57.8	70.5	66.6 (+ 8.8)	82.7 (+ 12.2)	78
N300	30.0	493.2	58.8	71.2	68.1 (+ 9.3)	82.1 (+ 10.9)	136
Avg	—	—	50.6	69.6	57.22 (+ 6.62)	80.84 (+ 11.24)	—

Table 2. Thermal impact of leakage power in 3D case.

Circuit	Area	HPWL	Max temperature			Time (s)
			With FLP	With TDLP	Error	
Ami33	39.2	21.1	136.2	212.2 (+ 76.0)	+ 55.8%	59
Ami49	1392.9	411.2	128.0	204.8 (+ 76.8)	+ 60%	68
N100	5.4	90.9	117.3	175.4 (+ 58.1)	+ 49.5%	77
N200	6.1	164.6	123.3	187.9 (+ 64.6)	52.4%	321
N300	8.8	224.6	135.7	198.6 (+ 62.9)	46.3%	450
Avg	—	—	128.1	195.78 (+ 67.68)	52.8%	—

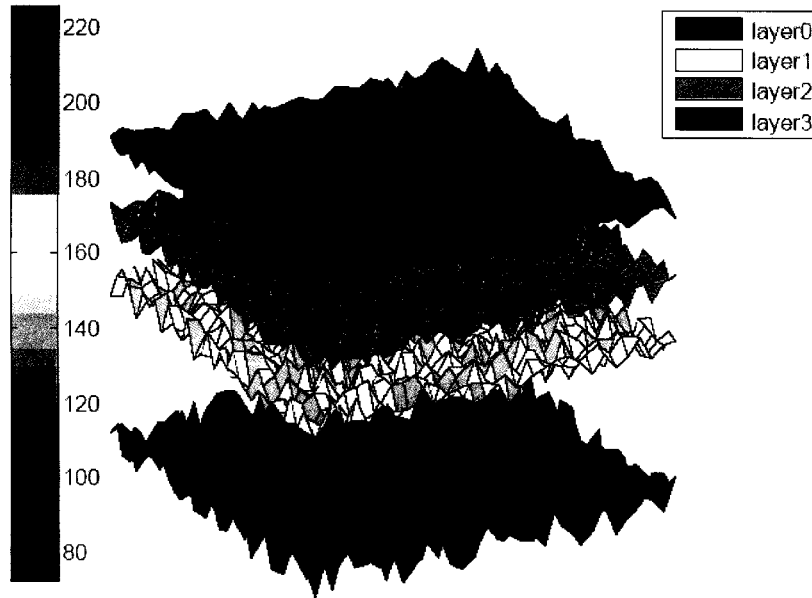


Fig. 7. The temperature distribution for n300 on four layers.

differences between layers are extreme high. As shown in Fig. 7, the ratio of leakage power varies greatly for different layers. Without the consideration of dependence of leakage power on temperature in 3D floorplanning, the temperature estimation can have about 50% error as shown in Table 2.

### 5.2. Impact of thermal-driven floorplanning with temperature-dependent leakage power

This section shows the effectiveness of our thermally optimized floorplanning flow in reducing chip temperature and leakage power as described in Sec. 4.2. In our experiments, we compare the results of the following three kinds of floorplanner: TF — traditional floorplanner which only optimize chip packing area and wirelength; TDF-FLP — thermal driven floorplanner which treat leakage power as fixed numbers during optimization process; TDF-TDLP — thermal driven floorplanner with temperature dependent leakage power considered. Figure 8 shows four-layer packing of n100 generated by TDF-FLP with the corresponding temperature distribution. Table 2 shows that without the thermal driven floorplanning, the maximal temperature can reach as high as 175°C. While in the packing generated by thermal driven floorplanning with consideration of leakage power (TDF-TDLP), the maximal temperature can be reduced to 128°C while the area and wirelength are remained without serious degradation.

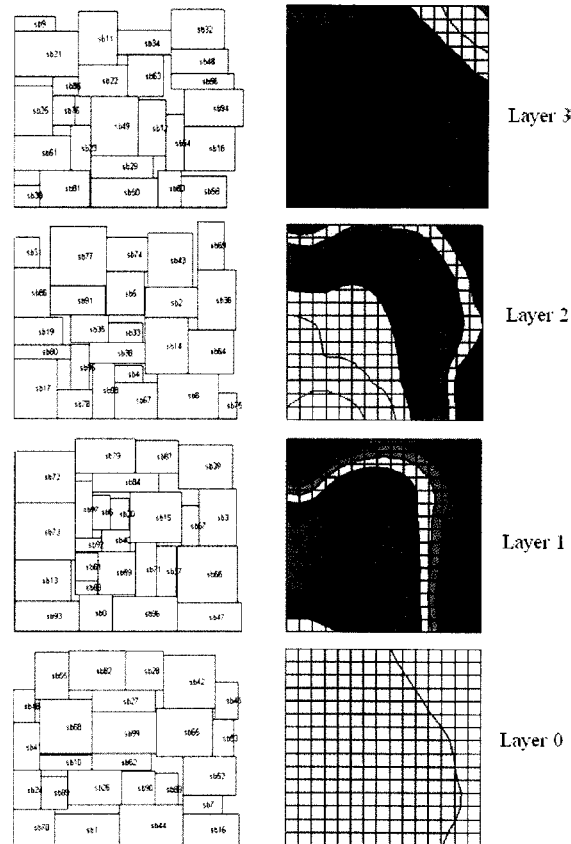


Fig. 8. Four-layer packing of n100 and the corresponding temperature distribution on each layer using TDF-TDLP.

Table 3. Impact of thermal-driven floorplanning with leakage power in 2D case.

Circuit	TF				TDF-FLP				TDF-TDLP									
	Area	HPWL	Avg T	Max T	LP%	Time	Area	HPWL	Avg T	Max T	LP%	Time	Area	HPWL	Avg T	Max T	LP%	Time
Ami33	123.2	58.3	49.7	81.6	17.43	4	126.7	62.7	49.6	75.9	13.54	67	126.7	63.75	49.5	73.8	11.86	108
Ami49	3663.0	861.6	41.7	77.8	15.79	10	3904.5	891.0	40.4	74.7	12.13	84	3904.5	871.0	39.8	72.4	10.40	136
N100	18.8	191.0	60.0	80.0	20.69	24	19.6	205.8	58.2	75.2	13.45	136	19.4	196.3	57.1	72.8	11.20	258
N200	19.4	367.1	66.6	82.7	20.84	78	19.6	375.4	62.8	76.1	14.01	195	19.5	387.1	61.8	73.4	12.07	391
N300	30.0	493.2	68.1	82.1	21.08	136	30.2	497.5	63.4	77.3	13.76	272	30.2	502.1	61.4	74.2	10.06	556
Avg	1	1	1	1	19.17	1	1.03	1.04	0.96	0.94	13.38	7.06	1.02	1.04	0.94	0.90	11.12	12.09

Table 4. Impact of thermal-driven floorplanning with leakage power in 3D case.

Circuit	TF				TDF-FLP				TDF-TDLP						
	Area	HPWL	Max T	LP%	Time	Area	HPWL	Max T	LP%	Time	Area	HPWL	Max T	LP%	Time
Ami33	39.2	21.1	212.2	36.54	59	39.7	23.2	170.9	25.65	279	40.1	24.3	146.2	21.75	787
Ami49	1392.9	411.2	204.8	38.30	68	1557.2	454.0	168.1	25.32	352	1462.5	467.5	143.5	20.93	895
N100	5.4	90.9	175.4	40.53	77	6.1	95.2	145.5	26.81	548	6.1	93.7	128.1	20.90	1349
N200	6.1	164.6	187.9	36.91	321	6.2	164.1	154.3	25.96	696	6.3	170.6	138.6	20.88	1622
N300	8.8	224.6	198.6	35.60	450	9.5	228.4	162.4	26.97	922	9.4	230.4	141.4	21.94	1933
Avg	1	1	1	37.58	1	1.07	1.05	0.82	26.14	4.25	1.06	1.07	0.71	21.28	10.67

Tables 3 and 4, respectively, show the results produced by the above three kinds of floorplanner in 2D and 3D cases. In these two tables, the columns AvgT and MaxT represent the average and max temperatures of the whole chip, and the column LP% shows the proportion of leakage power to total power (total power = leakage power + dynamic power). *All the temperature values in these two tables are measured with temperature-dependent leakage power.* From Table 3, we can see that, on average, TDF-FLP can reduce average and peak temperatures by 4% and 6% with slight increase in packing area and wirelength. Compared with TF, TDF-FLP has 5.8% less LP% results on average. Besides, TDF-FLP needs nearly 6 times more CPU time to do thermal analysis. Compared with TDF-FLP, TDF-TDLP can further optimize average and max temperatures by 2% and 4%, and reduce LP% to 11.12% with similar packing area and wirelength. TDF-TDLP needs another more 5 times of CPU time to do leakage aware iterative thermal analysis. Table 4 shows the results for 3D case. Compared with TF, TDF-FLP optimizes peak temperature by 18% and reduces LP% from 37.58% to 26.14% with 7% and 5% increases in packing area and wirelength. However, TDF-TDLP outperforms TDF-FLP by 11% in peak temperature and nearly 5% lower LP% with about double CPU time.

## 6. Conclusion

In this paper, we investigate the impact of leakage power on thermal profile in floorplanning stage and propose an efficient thermal optimized floorplanning flow with leakage power considered to reduce chip peak temperature and leakage power in 2D and 3D floorplanning. Without the consideration of dependence of leakage power on temperature in 3D floorplanning, the temperature estimation can have about 50% error. Experiment results show that thermal-driven optimization algorithm is efficient in reducing chip temperature and saving leakage power consumption.

## References

1. S. Nassif *et al.*, Full chip leakage estimation considering power supply and temperature variations, *Proc. Int. Symp. Low Power Electronics and Design*, August 2003, pp. 78–83.
2. Semiconductor Industry Association, International technology roadmap for semiconductors (2005), <http://public.itrs.net>.
3. A. Chandrakasan, W. Bowhill and F. Fox, *Design of High-Performance Microprocessor Circuits* (IEEE Press, 2001).
4. A. Gupta *et al.*, STEFAL: A system level temperature-and floorplan-aware leakage power estimator for SoCs, *Proc. VLSI Design* (2007), pp. 559–564.
5. W. Liao *et al.*, Microarchitecture level power and thermal simulation considering temperature dependent leakage model, *Proc. Int. Symp. Low Power Electronics and Design* (2003), pp. 211–216.
6. W. Huang *et al.*, The need for a full-chip and package thermal model for thermally optimized IC designs, *Proc. Int. Symp. Low Power Electronics and Design*, August 2005, pp. 245–250.

7. K. Sankaranarayanan *et al.*, A case for thermal-aware floorplanning at the micro-architectural level, *J Instruction Level Parallelism* **8** (2005) 8–16.
8. K. Sankaranarayanan *et al.*, Microarchitectural floorplanning for thermal management: A technical report, University of Virginia (2005).
9. Y. Han *et al.*, Temperature aware floorplanning, *Proc. Temperature-Aware Computer Systems*, June 2005.
10. Healy *et al.*, Microarchitectural floorplanning under performance and thermal tradeoff, *Proc. Conf. Design, Automation and Test in Europe* (2006), pp. 1288–1293.
11. W. L. Hung *et al.*, Thermal-aware floorplanning using genetic algorithms, *Proc. 6th Int. Symp. Quality of Electronic Design* (2005), pp. 634–639.
12. Y. Cheon *et al.*, Power-aware placement, *Proc. Conf. Design Automation* (2005), pp. 795–800.
13. B. Obermeier *et al.*, Temperature aware global placement, *Proc. 2004 Conf. Asia South Pacific Design Automation* (2004), pp. 143–148.
14. M. Ekpanyapong *et al.*, Thermal aware 3-D microarchitectural floorplanning, Technical Reports, Georgia Tech., December 2004.
15. J. Cong *et al.*, A thermal-driven floorplanning algorithm for 3D ICs, *Proc. 2004 IEEE/ACM Int. Conf. Computer-Aided Design* (2004), pp. 306–313.
16. W.-L. Hung *et al.*, Interconnect and thermal-aware floorplanning for 3D microprocessors, *Proc. 7th Int. Symp. Quality Electronic Design* (2006), pp. 98–104.
17. Z. Li *et al.*, Hierarchical 3-D floorplanning algorithm for wirelength optimization, *IEEE Trans. Circuits Syst. I* **53** (2007) 2637–2646.
18. B. Goplen and S. Sapatnekar, Efficient thermal placement of standard cells in 3D ICs using a force directed approach, *Proc. Int. Conf. Computer-Aided Design* (2003), pp. 86–89.
19. G. S. Ohm, *The Galvanic Circuit Investigated Mathematically* (D. Van Nostrand Co., Kraus, New York, 1891).
20. Y. Yang, Z. P. Gu, C. Zhu, R. P. Dick and L. Shang, ISAC: Integrated space and time adaptive chip-package thermal analysis, *IEEE Trans. CAD* **26** (2007) 86–99.
21. H. Eisenmann and F. M. Johannes, Generic global placement and floorplanning, *Proc. Conf. Design Automation* (1998), pp. 269–274.
22. K. Vorwerk, A. Kennings and A. Vannelli, Engineering details of a stable analytic placer, *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, November 2004, pp. 573–580.