Thermal Effects with Leakage Power Considered in 2D/3D Floorplanning

Pingqiang Zhou Yuchun Ma Qiang Zhou Xianlong Hong
Tsinghua University, Beijing, P.R.China
E-mail: myc@mail.tsinghua.edu.cn

Abstract

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 2D and 3D floorplanning. Our results show that chip temperature can increase by about 11℃ in 2D design and 68℃ 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℃ 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℃ temperature reduction and 16.3% less leakage power proportion.

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 (the main component of leakage current) negligible. 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 generation of CMOS technologies, the number is expected to rise to 50% for the next generation technologies [1][2].

As indicated in [3], the sub-threshold current of a transistor has a super-linear dependency on temperature. For newer 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. Aseem Gupta et al in [4] show that different floorplans have different thermal profiles, thus have different leakage power.

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.

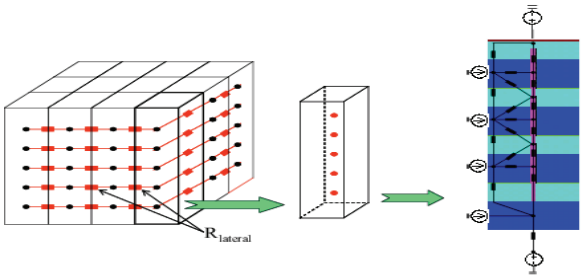
The rest of the paper is organized as follows. In section 2, we review the related work in leakage power modeling and thermal-aware floorplanning/placement. In section 3, we present the thermal and leakage power models. Section 4 formulates our thermal-driven floorplanning flow with temperature-dependent leakage power considered. Section 5 shows the experimental results and the conclusions are provided in section 6.

2. Related work

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

A. Chandrakasan et al. in [3] show that the subthreshold current for a transistor is a super exponential function of its supply voltage, threshold voltage and temperature. S. Nassif et al. [5] propose a model to estimate the leakage power for a chip while considering power supply and temperature variations. W. Liao et al. [6] also propose a temperature-dependent leakage power model with an assumption that the whole chip is at one uniform temperature. W. Huang et al in [7] uses 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 [8] [9]. Han et al. in [10] conduct a study on the floorplanner's effectiveness in lowering the maximum processor temperatures. Recently, Healy et al. [11] have proposed thermal-driven and communication profile-driven floorplanning to reduce peak temperatures. W-L. Hung et al [12] use



(a) Tiles Stack Array (b) Single Tile Stack (c)Tile Stack Analysis

Fig. 1 Resistive thermal model for a 3D IC

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. [13] and Obermeier et al. [14] respectively.

Besides, Ekpanyapong et al. [15] and Cong et al. [16] have proposed thermal-driven and communication profile-driven 3D floorplanning. Hung et al. propose thermal-aware floorplanning for 3D microprocessors [17], the power consumption of interconnect is considered during floorplanning. Recently, Li et al. developed a hierarchical 3D floorplanning algorithm [18]. Goplen and Sapatnekar [19] developed force-directed thermal-aware standard-cell placement algorithms. These techniques demonstrate good and scalable performance.

However, though previous floorplanning algorithms 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.

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 [20]. 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.

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

where A is an $N \times N$ sparse thermal conductivity matrix. T and $P(t)$ are $N \times 1$ temperature and power vectors. Li etc. 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 loss of accuracy [21]. 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

$$P_{leakage\ i} = A_i \times \alpha \times e^{\beta \times (T_i - T_{base})} \qquad (2)$$

where A_i and T_i are the area and temperature of the cell. α and β are empirical factors that have different values for different technologies(Typical values found in [9] are $\alpha = 1\times10^5$ W/m^2 and $\beta = 0.025$ for 130nm). T_{base} is the reference temperature at which α and β are defined. Fig. 2 shows the relation between temperature and leakage power. The leakage power increases with the temperature super-linearly.

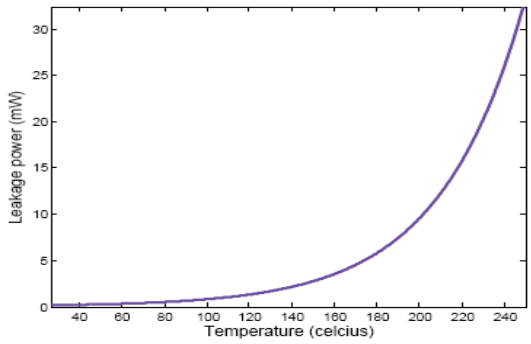


Fig.2 Relation between leakage power and temperature

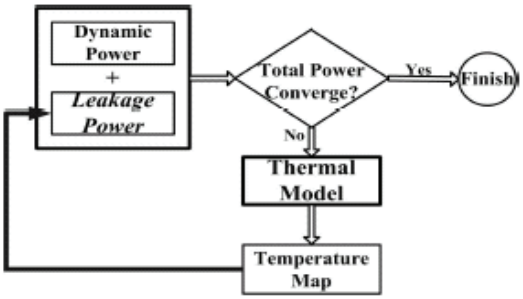


Fig. 3 Full-chip thermal model closes the loop for accurate leakage power calculation

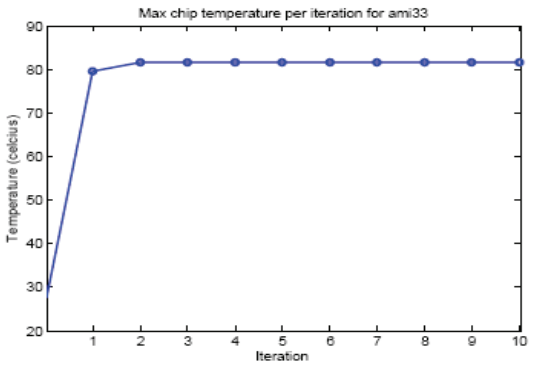


Fig. 4 Iteration process in thermal analysis

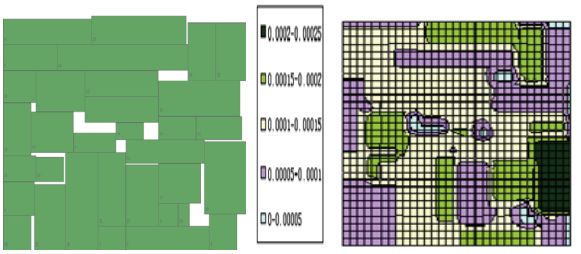
4. Thermal-driven Floorplanning with Leakage Power

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 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.

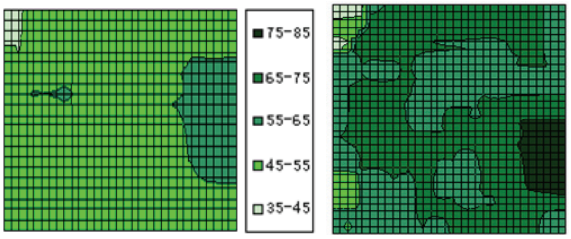
4.1 Leakage-aware iterative thermal analysis

As shown in section 3, leakage power and temperature value 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 section 3.2. The total power distribution will be updated with the new leakage power. And the thermal distribution can be further updated using the thermal model in section 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 a test case 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. Fig. 5 (a) shows the corresponding dynamic power distribution. Then we use the thermal tool introduced in section 3 to get the thermal profiles for this packing without and with leakage power consideration, as shown in Fig. 5 (c) and Fig. 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 centigrade on average.



(a) A 2D packing for circuit ami33 (b) Dynamic Power distribution



(c) Thermal Profile with no leakage (d) Thermal Profile with leakage power

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

4.2 Thermally Optimized Floorplanning flow

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 but also take 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 [22][23] and extend it to deal with 3D thermal driven floorplannig problem. We choose force-directed placement engine because of its good performance and scalable feature.

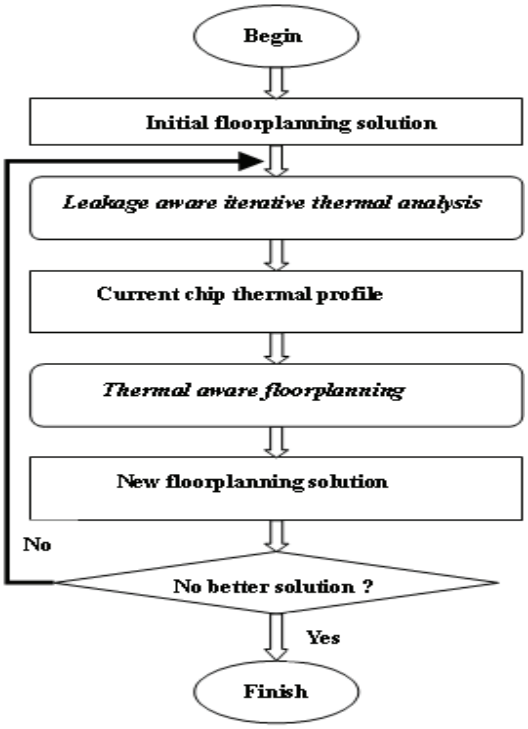


Fig. 6 Thermally optimized floorplanning flow

5. Experiment Results

Our algorithm is implemented in C++, and the experiments are performed on a workstation with 3.0 GHz CPU and 4GB physical memory. We use five typical MCNC and GSRC benchmarks [15] in our experiments. We assign a power density of random number between 1.0×10^5 (W/m^2) and 5.0×10^6 (W/m^2) 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 results are Area in 1.0×10^4 μm^2, Wirelength (HPWL) in mm, Temperature in ℃, and Time in seconds (s).

5.1 Thermal Impact of leakage power

Table 1 and table 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 centigrade obtained by the model described in section 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 centigrade when we take temperature-dependent leakage power into account. In 3D floorplanning, however, the max temperature increases by 67.7 centigrade, which is about 5 times worse than the number in 2D case. Without the consideration of leakage power in 3D floorplanning, the temperature estimation can have about 50% error.

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 section 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 wirelengh; *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.

Table 3 and Table 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

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

Table 2 Thermal Impact of leakage power in 3D case

Circuit	Area	HPWL	Max Temperature		Time
			With FLP	With TDLP	
Ami33	39.2	21.1	136.2	212.2 (+76.0)	59
Ami49	1392.9	411.2	128.0	204.8 (+76.8)	68
N100	5.4	90.9	117.3	175.4 (+58.1)	77
N200	6.1	164.6	123.3	187.9 (+64.6)	321
N300	8.8	224.6	135.7	198.6 (+62.9)	450

(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. As far as we know, this is the first paper which optimizes thermal effects with leakage power considered in 3D floorplanning. Experiment results show that our optimization algorithm is efficient in reducing chip temperature and saving leakage power consumption.

7. References

[1] H. Su, S. Nassif et al.. "Full Chip Leakage Estimation Considering Power Supply and Temperature Variations". In Proc. Int. Symp. Low Power Electronics. & Design, Aug. 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, *De sign of High-Performance Microprocessor Circuits*. IEEE Press, 2001.

[4] Aseem Gupta et al.. "STEFAL: A System Level Temperature-And Floorplan-Aware Leakage Power Estimator for SoCs". In Proc. VLSI Design 2007: pp.559-564.

[5] W. Liao et al., "Microarchitecture Level Power and Thermal Simulation Considering Temperature Dependent Leakage Model," in Proceedings ofInternational Symposium on Low Power Electronics and Design 2003, pages 211-216.

[6] W. Huang et al, "The need for a Full-Chip and Package Thermal Model for Thermally Optimized IC Designs", in Proc. Int. Symp. Low Power Electron. Design, Aug. 2005,. pp. 245–250

[7] K. Sankaranarayanan et al., "A Case for Thermal-Aware Floorplanning at the Microarchitectural Level," *Journal of Instruction Level Parallelism*, vol. 8, pp. 8--16, Oct. 2005

[8] K. Sankaranarayanan et al., "Microarchitectural Floorplanning for Thermal Management: A Technical Report," *University of Virginia*, 2005.

[9] Y. Han et al., "Temperature Aware Floorplan ning," in Proc. Temperature-Aware Computer Systems, Jun. 2005.

[10] Healy, Ekpanyapong et al., "Microarchitectural Floorplanning Under Performance and Thermal

Tradeoff," Proceedings of the conference on Design, automation and test in Europe 2006, pp. 1288 - 1293.
[11] W.L. Hung et al.,"Thermal-Aware Floorplanning Using Genetic Algorithms," Proceedings of the 6th International Symposium on Quality of Electronic Design, 2005 pp. 634 - 639 .
[12] Y. Cheon et al., "Power-Aware Placement," Proceedings of the conference on Design automation, 795 – 800, 2005.
[13] B. Obermeier et al., "Temperature Aware Global Placement," Proceedings of the 2004 con ference on Asia South Pacific design automation, 2004. 143 - 148
[14] M. Ekpanyapong et al., "Thermal Aware 3-D Microarchitectural Floorplanning," *Technical Reports, Georgia Tech*, December 2004.
[15] J. Cong, J. Wei et al.,"A Thermal-Driven Floorplanning Algorithm for 3D ICs", Proceedings of the 2004 IEEE/ACM International conference on Computer-aided design, *2004*, pp. 306 - 313
[16]W.-L. Hung, et al., "Interconnect and thermal-aware floorplanning for 3D microprocessors," Proceedings of the 7th International Symposium on Quality Electronic Design. 2006. pp. 98 - 104
[17] Z. Li, et al., "Hierarchical 3-D floorplanning algorithm For wirelength optimization," *IEEE Trans. on Circuits and Systems I*, vol.53, issue.12, pp. 2637-2646, 2007.
[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, pp. 86-89, Nov. 2003.
[19] G. S. Ohm, "The Galvanic circuit investigated mathematically," New York, D. Van Nostrand Co., Kraus, 1891.
[20] Yonghong Yang and Zhenyu (Peter) Gu and Changyun Zhu and Robert P. Dick and Li Shang, "ISAC: Integrated Space and Time Adaptive Chip-Package Thermal Analysis" *IEEE Trans. on CAD*. Volume: 26, Issue: 1, 86-99 2007.
[21] H.Eisenmann and F.M.Johannes, "Generic Global Placement and Floorplanning", in Proc. the conference on Design automation, pages.269-274, 1998
[22] K. Vorwerk, A. Kennings, and A. Vannelli,"Engineering details of a stable analytic placer," in Proc. of IEEE/ACM International conference on Computer-aided design, pp. 573- 580 Nov.2004.

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

Circuit	TF						TDF-FLP						TDF-TDLP					
	Area	HPWL	Avg T	MaxT	LP%	Time	Area	HPWL	Avg T	MaxT	LP%	Time	Area	HPWL	Avg T	MaxT	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	MaxT	LP%	Time	Area	HPWL	MaxT	LP%	Time	Area	HPWL	MaxT	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