Research Journal of Applied Sciences, Engineering and Technology 8(10): 1279-1284, 2014
DOI:10.19026/rjaset.8.1096
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: July 24, 2014 Accepted: September 13, 2014 Published: September 15, 2014

Research Article

Thermal Aware Floor planning Technique for Nano Circuits

G. Nallathambi, D. Gracia Nirmala Rani and S. Rajaram
Department of ECE, Thiagarajar College of Engineering, Madurai-625015, India

Abstract: The strongest challenge that a VLSI designer has to face today is the extremely high heat generation within a chip which not only degrades the performance but also the yield and reliability are greatly affected. The situation even became worse with the more number of wires in a single chip and due to path to path temperature variations within the chip. Nowadays VLSI circuits have immense variations in temperature and their linear relationship between metal resistance and temperature causes distinct delay through wires of the same length. The aim of this study is to analyze floor planning algorithms and wire plan methods to reduce the temperature dependent delay in global wires. We propose a temperature dependent wire delay estimation method for thermal aware floor planning algorithms, which takes into account the thermal effect on wire delay. The experiment results show that a shorter delay can be achieved, the congestion and reliability issues as they are closely related to routing and temperature using the proposed method.

Keywords: Congestion, floor planning, thermal, wire delay

INTRODUCTION

With decreasing feature sizes and increasing transistor counts, power density in VLSI circuits has increased dramatically. Since the heat generated by a VLSI circuit is proportional to its power density, the corresponding rise in on-chip temperatures adversely impacts reliability, circuit performance and cooling costs. According to ITRS, thermal management is projected as one of the most challenging issues in the design of future high-performance integrated circuits (Elmore, 1948). Different functional units in a circuit can have different switching activity rates, leading to uneven power dissipation among the various routing path on the chip. Due to low thermal conductivity of silicon, the rate of lateral heat propagation in a chip is slow, causing localized heating to occur much faster than chip-wide heating. This can result in an un even temperature distribution on the chip, creating on-chip thermal gradients and "thermal hot spots" caused by localized areas of high power densities.

Wire delay is temperature dependent since resistivity in metal increases with temperature. Global wires are routed over many functional blocks which exhibit a large thermal gradient along the length of the wire; this non uniform thermal problem can significantly degrade wire performance. In nanometer technologies more and more circuits are limited by wire delay rather than gate delay, therefore optimization of wire delay is very important and sometimes crucial. Thermal aware global routing algorithms for enhancing reliability are also discussed. Global signal wires work has been done on thermal aware floor planning, electrical resistivity in wires is constant and thermal gradients in the substrate has no slam on wire delay is surmised by all. This posits is in general invalid and increasingly erroneous in nanometer high performance designs where immense temperature gradients already subsist in the substrate (Ajami *et al*., 2005).

The problem of estimating the temperature dependent wire delay during the floor planning stage is studied in this study (Black, 1969). The impact of non uniform thermal profile on the delay in wires is summarized. A new way to estimate the wire delay in thermal aware floor planning algorithms has been proposed. The proposed algorithm takes the delay, rather the wire length, excessive delay increase caused by high temperature is reduced is taken as one of the optimization goals. In addition, we also consider the impact of routing congestion and the reliability of wires, which are prominent metrics in analyzing floor plans in a realistic setting.

The aim of this study is to analyze floor plan algorithms and wire plan methods to reduce the incrementing temperature dependent delay in global wires.

LITERATURE SURVEY

The hot spot tool as proposed by Skadron *et al*. (2004) is a computationally effective and easy to use thermal modeling tool to estimate the thermal effect as block level. It gives a simple compact model to take in to account the heat dissipation with in each functional

Corresponding Author: G. Nallathambi, Department of ECE, Thiagarajar College of Engineering, Madurai-625015, India

block. Basic idea of the scheme is that if we know the thermal resistance and power distribution of a given floor plan then we can calculate the temperature of each block:

$$P_j = R^t \times T_j \quad (1)$$

The original compact thermal model has been simplified in the proposed scheme for partition-driven placement which allows using temperature in the inner placement loop as a constraint for a better thermal distribution.

Han *et al*. (2005) the heat diffusion model is first proposed by Han *et al*. (2005). According to this model, heat diffusion between any two adjacent blocks is proportional to their temperature difference and length of the shared boundary between them. As power density is directly proportional to temperature difference is equivalent to power density difference. Thus heat diffusion between two blocks can be expressed as:

$$H(d_{1,}d_2) = (d_1 - d_2) \times L \quad (2)$$

L is the shared length whereas the total heat diffusion to be:

$$H(d) = \sum H(d, d_i) \quad (3)$$

The goal of a good floor plan is to minimize the chip area, make the subsequent routing phase easy and to minimize the wire length and hence the cost. Huang *et al*. (2006) proposed a genetic algorithm based thermal aware floor planning that aims at reducing hotspots and distributing the temperature uniformly across the chip. It also takes into account the traditional design goal, chip area.

Jaffari and Anis (2007) has given a thermal aware placement technique for FPGAs to reduce maximum temperature and on chip temperature gradients. A new cost function has been given for the simulated annealing core of the placement tool which is based on electrostatic charge model instead of extracting thermal profile at each simulation run and claimed to achieve better results with an algorithmic complexity linear with the number of logic blocks. Liu *et al*. (2008) presented an efficient and effective simultaneous hot spot avoid embedding and Thermal aware Routing (TMST) method, where hot spot embedding avoid tree topology located in area with high temperature possibility and thermal aware routing reduce skew in tree path with more smooth temperature area.

Ning and Zhonghua (2008) present a Gauss scidal method of floor planning for thermal aware design where they proposed an incremental iterative method to solve the thermal model of hierarchical floor planning which can avoid hot spots in the design of chip. Schafer and Kim (2008) proposed a thermal aware design and hotspot reduction technique from the gate level net list. Han demonstrated that how different floor plans can affect the maximum temperature of the chip. The temperature difference of different floor plans of the same design can be as high as 30 degree centigrade. The proposed model includes the traditional goal of area and wire length optimization along with the heat diffusion measure as an approximation to temperature and solved the problem using simulated annealing algorithm.

PROPOSED METHODOLOGY FOR THERMAL AWARE FLOOR PLANNING

Hot Spot tool has been developed by University of Virginia. Hot Spot is an accurate and fast thermal modeling tool suitable for use in architectural studies. For every input floor plan, Hot Spot generates an equivalent circuit of thermal resistances and capacitances. The equivalent circuit of thermal resistances and capacitances corresponds to micro architecture blocks and essential aspects of the thermal package. Hot Spot has a simple set of interfaces. The main advantage of Hot Spot is its compatibility with the kinds of power/performance models used in the computer-architecture community. It does not require detailed design or synthesis description. It makes it possible to study thermal evolution over long periods of real, full-length applications.

Floor plan, the algorithm invokes routines in Hot Spot to compute the temperature distribution and uses the maximum temperature as one of the metrics in the cost function. The other metrics in the cost function are total area and total wire length. The connectivity information in Hot Floorplan is stored in a two-dimensional connectivity matrix and manhattan distance is used to estimate the wire length between two endpoints in a wire (Chen and Chang, 2006).

The power density of today's high performance IC chips is too large and as a result, identification of hot spots and subsequently design of cooling mechanisms including thermal vias and heat sinks has become very important. For analysis of hot spots, heat diffusion equations are usually employed to calculate their effects. In thesis, we have described a completely new model to analyze the behavior of hot spots based on a purely geometric approach.

A geometric model is proposed to evaluate the cumulative thermal power at any point of the chip floor. This is followed by a discrete version of the model that works in the grid domain.

Then a simulation based approach is presented that identifies the hot spots and Zone on chip floor using the above models. Contrary to natural intuition, it is found

that some non source points on the chip floor can become more heated than heat generating source points on the chip floor can become more heated than heat generating source points due to cumulative heating effect of nearby source. To improve over the time consuming simulation based approach, a voronoi diagram based technique is further proposed that helps in quickly identifying the hot zones. To get a more accurate analysis, the simulation based approach can then be administered in and around these hot zones. The final outcome of this approach is an overall saving in simulation time. Further, we describe a procedure to dilate the hot zones by shifting some of the heat generating sources away from each other in order to make the hot zones cool down below the critical level. The movements are made within some restrictions so that the topology of the voronoi diagram constructed around the point heat sources does not change. This technique while dilating the heat sources will not change the floor plan adjacency relations significantly.

We introduce a new concept of density in an ensemble of points on a 2D floor and prove several geometric results with maximum/minimum density of source points on the chip floor. In the light of these results, the algorithms required to identify the densest (hot) zone, as well as the most rarefied (cool) zone turn out to be quit simple. The concept of density has then been compared with an existing concept of discrepancy and it is shown that there appears to be some distinct advantage in applying the concept of density over discrepancy in thermal analysis of hot spots in a chip.

The delay of global wires subject to large temperature variations is no longer linearly proportional to wire length. To have an accurate estimation of wire delay, the temperature effect including thermal gradients has to be considered. Since the thermal profile on the die is mainly determined by the locations of macro blocks, it is, therefore, possible to perform temperature dependent delay estimation at the floor planning stage. Our wire planning method is described in Algorithm 1. Given a floor plan, together with the associated connectivity matrix and the thermal map, we compute the total delay, the maximum congestion and the average reliability of all wires. The layout of each wire is determined by performing L-shape routing between the centers of the connecting blocks. Once the physical layout of the wire is known, we record the blocks over which the wire is routed. The temperature profile along the wire and the wire delay are then calculated using (1) and (4). With thermal effects taken in account, the two paths in the bounding box of two end points of a wire can have different delay although their length are the same. In our algorithm, we choose the path with a shorter delay, which is different from Hot Floor plan where the two paths are considered as identical. The temperature profile is also used to evaluate the wire reliability in terms of Mean Time to Failure (MTTF). In addition, a congestion map made up of a two dimensional matrix is updated with the route of the wire to evaluate the rout ability of the floor plan.

TEMPERATURE ESTIMATION AND WIREDELAY CALCULATION

Detailed routing information of an interconnect is unknown at the Foor planning stage and Wire length is usually estimated as the Manhattan distance between two connected blocks (Chakraborty *et al*., 2008). The half perimeter can be along the lower bend or the upper bend and when thermal effects are not considered, the resistance is constant along the interconnect meaning that either route would have the same delay. For thermal aware wire delay estimation, however, different routes of the same length can have Very different delay since they are subject to different thermal problems.

The initial stage of physical implementation of VLSI circuits is Floor planning that influences the quality of the final design. The main design tasks incorporates macro block placement, global wire planning and Power/Ground network design in the floor planning stage. The total area and wire length is the only thing optimized by Traditional floor planning algorithms. Thermal aware floor planning algorithms were called so, when a thermal issue becomes more important, the maximum temperature is also pre pended to the cost functions as discussed in Sankaranarayanan *et al*. (2005) and Gupta *et al*. (2007). The temperature distribution in the whole chip is disturbed due to the thermal coupling between high power consumption blocks; thermal aware floor planning algorithms takes peak temperature along with area and wire length in the evaluation of a floor plan. By using compact thermal models, which can evaluate the temperature profile in a very efficient way the peak temperature is estimated in Tsai and Kang (2000). A very representative thermal aware floor planner, Hot floor plan is the floor planning tool which proposed in Sankaranarayanan *et al*. (2005). The topology of the floor plan is symbolized in Normalized Polish Expression and the optimization process is implemented as a simulated annealing process. Each and every candidate floor plan, the algorithm refers to routines in Hot Spot (Cho *et al*., 2005) to calculate the temperature distribution and uses the maximum temperature as one of the metrics in the cost function. The total area and total wire length were the other metrics in the cost function. The connectivity information in Hot Floor plan is saved in a two-dimensional connectivity matrix and uses Manhattan distance to estimate the wire length between two endpoints in a wire (Lu and Pan, 2009).

Heat diffusion from the substrate as well as self heating causes high temperature in global wires. The temperature within the interconnect for a given substrate temperature can be expressed as (Gupta *et al*., 2008):

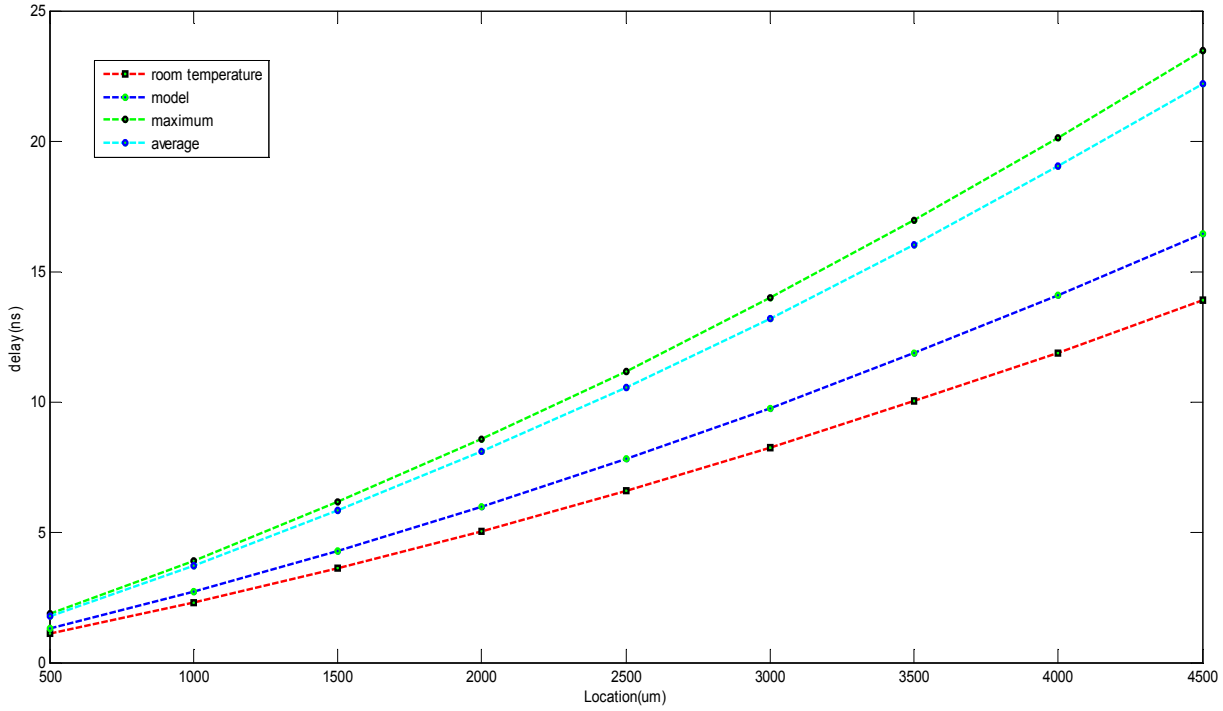

Fig. 1: Delay estimation in thermal profiles

$$T(x) = T_{sub} + \frac{\theta}{\lambda^2}[1 - \frac{\sinh \lambda x + \sinh \lambda (L - x)}{\sinh \lambda x}],$$
$$0 \le x \le L \quad (4)$$

where,

T_{sub} = Temperature of the substrate
$\theta \& \lambda$ = Constants depending on the chosen metal layer
L = Length of the wire

The length (L) of interconnects which are larger than the heat diffusion length whose peak temperature rise is equal to $\frac{\theta}{\lambda^2}$:

$$\lambda^2 = \frac{1}{k_m}[\frac{k_{ins}^*}{t_m t_{ins}} - \frac{I_{rms}^2 \rho_i \beta}{\omega^2 t_m^2}] \quad (5)$$

$k_m \& k_{ins}^*$ = Thermal conductivity of metal and insulator
$k_m \& t_{ins}$ = Thickness of metal and insulator
I_{rms} = Average current density in interconnect

There will be high temperature rise in global wires because the distance from the substrate is larger than local wires. The linear relationship between the electrical resistance of metal and its temperature can be expressed as:

$$R(x) = R_0(1 + \beta T(x)), 0 \le x \le L \quad (6)$$

where, R0 is the resistance at reference temperature, β is the temperature coefficient, T(x) is the temperature profile along the wire (Eq. 6).

For copper the value of β at room temperature is 3.9*10-3, which means for every 10°C rise in temperature, the resistance increases by 3.9%. Signal propagation delay through the interconnect of length L as in the distributed RC Elmore (1948) delay model (Gupta *et al*., 2008), can be written as:

$$D = R_d(C_L + \int_0^L C_0(x)dx) + \int_0^L r_0(x)dx \times (\int_x^L c_0(\tau)d\tau)dx \quad (7)$$

Which is different from Hot Floor plan where the two paths are considered as identical? The temperature profile is also used to evaluate the wire reliability in terms of Mean Time to Failure (MTTF). In addition, a congestion map made up of a two dimensional matrix is updated with the route of the wire to evaluate the rout ability of the floor plan. The congestion map is useful because in our algorithm more wires are likely to be routed in regions with a low temperature, potentially causing routing congestion. We now describe the congestion map and the reliability metrics used in the algorithm (Fig. 1).

CONGESTION MAP

Our congestion map is made up of a two dimensional matrix, which divides the floor plan into a congestion grid. The size of the routing cell in the grid is 64 μm×64 μm. During the routing of wires, the value

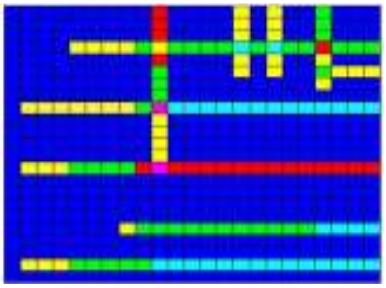

Fig. 2: Congestion map-apte circuit

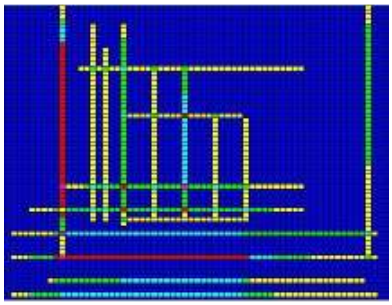

Fig. 3: Congestion map-Hp circuit

of a cell is incremented if a wire passes through the cell. A higher value means more wires passes through that cell. An example congestion map is shown in Fig. 2, where the congestion cells are plotted in colors over the floor plan. Cells in red color indicate congestion hotspots. The maximum value in the congestion matrix can be used to evaluate the rout ability of the design. This result occurs because the thermal-aware floor planner places units so that the heat flow in the chip is to be maximized allowing hotter units to cool down faster by placing them close to cooler ones. This can be a good way to flatten the overall chip temperature (at a coarse parts in a unit might get hotter than others due to the difference of intrinsic switching activity of their transistors). A congestion map made up of a two dimensional matrix is updated with the route of the wire to evaluate the routing ability of the floor plan (Fig. 3).

RESULTS AND DISCUSSION

The thermal aware Routing tool was run for the temperature part of the objective function (thermal aware routing). The results of the two cases were compared with each other in terms of three main parameters: area, wire length and temperature. The tool was run for multiple times for MCNC benchmarks like hp, apte.

From the results, the following points were inferred:

- Effect of Thermal Aware routing on Temperature is the maximum temperature of the chip reduces considerably when thermal aware routing is used, in comparison to normal routing or routing plan without temperature factor. The normalized reduction of maximum chip temperature is validated through various benchmark circuits. The percentage reduction of maximum chip temperature is noted.
- Effect of Thermal Aware Routing on Area of the Chip is validated with various benchmark circuit. As a result of thermal aware routing, there is a slight increase in chip area. Thus, the effect of thermal aware floor planning on chip area was found to be minimal in most of the benchmarks. Hence, the temperature reduction is a result of placement of blocks in a thermal aware manner rather than the increase in area.
- Effect of Thermal Aware Floor Routing on Wire Length shows a slight increase in wire length of the chip. The results are normalized with the increase in wire length in various benchmarks due to thermal aware routing. Thus, there is no significant increase in wire length as a result of thermal aware routing. In hp benchmark, sometimes a small reduction in wire length was observed. The floor plans in both the cases were plotted and it was noticed that the reduction in wire length was because of interconnects between them.
- Effect of Thermal Aware Floor Planning on the Placement of the Blocks is described with the cooling of the blocks in a floor plan happens due to lateral spreading of heat through silicon blocks. If a hot block is placed besides cooler blocks, lateral spreading of heat takes place. As a result, the temperature of the hot block is reduced. Thermal aware floor planning tool uses this concept of lateral heat diffusion.

CONCLUSION

In this study, we proposed a wire delay estimation method at the floor planning stage which takes into account the performance degradation in wires due to thermal effect. The method is implemented in Hot Floor plan and evaluated congestion using the MCNC benchmarks the experimental results show that in the presence of thermal gradients shorter wire length does not always produce shorter delay. The proposed method, on the other hand, can achieve a better total delay and HOT SPOT identification of IC circuit. The congestion maps give a visual impression of the congestion of the benchmarks and serve as confirmation that our implementation is correct. Finally, this study can serve as a stepping-stone towards a better understanding and application of thermal aware delay modeling at the floor planning stage. The goal of the study is to focus on the ongoing research in the field of thermal balanced design for nanometer ICs and to aware the challenges that are to be taken by the future VLSI designers.

REFERENCES

Ajami, A., K. Banerjee and M. Pedram, 2005. Modeling and analysis of nonuniform substrate temperature effects on global ULSI interconnects. IEEE T. Comput. Aid. D., 24(6): 849-861.

Black, J., 1969. Electro migration-a brief survey and some recent results. IEEE T. Electron. Dev., 16(4): 338-347.

Chakraborty, A., K. Duraisami, A. Sathanur and P. Sithambaram, 2008. Dynamic thermal clock skew compensation using tunable delay buffers. IEEE T. VLSI Syst., 16(6): 639-649.

Chen, T.C. and Y.W. Chang, 2006. Modern floorplanning based on b*-tree and fast simulated annealing. IEEE T. Comput. Aid. D., 25(4): 637-650.

Cho, M., S. Ahmedtt and D.Z. Pan, 2005. TACO: Temperature aware clock-tree optimization. Proceeding of IEEE/ACM International Conference on Computer-Aided Design (ICCAD, 2005), pp: 582-587.

Elmore, W.C., 1948. The transient response of damped linear networks with particular regard to wideband amplifiers. J. Appl. Phys., 19(1): 55-63.

Gupta, A., N. Dutt, F. Kurdahi, K. Khouri and M. Abadir, 2008. Thermal aware global routing of VLSI chips for enhanced reliability. Proceeding of 9th International Symposium on Quality Electronic Design, pp: 470-475.

Gupta, A., N. Dutt, F. Kurdahi, K. Khouri and M. Abadir, 2007. LEAF: A system level leakage-aware floorplanner for SoCs. Proceeding of the 2007Asia and South Pacific Design Automation Conference (ASP-DAC, 2007), pp: 274-279.

Han, Y., I. Koren and C.A. Moritz, 2005. Temperature aware floor planning. Proceeding of Workshop on Temperature Aware Microarchitectures.

Huang, W., S. Ghosh, K. Velusamy and K. Sankaranarayanan, 2006. HotSpot: A compact thermal modeling methodology for early-stage VLSI design. IEEE T. VLSI Syst., 14(5): 501-513.

Jaffari, J. and M. Anis, 2007. Thermal-aware placement for FPGAs using electrostatic charge model. Proceeding of 8th International Symposium on Quality Electronic Design (ISQED, 2007), pp: 666-671.

Liu, C., J. Su and Y. Shi, 2008. Temperature-aware clock tree synthesis considering spatiotemporal hot spot correlations. Proceeding of 26th IEEE International Conference on Computer Design, pp: 107-113.

Lu, K. and D. Pan, 2009. Reliability-aware global routing under thermal considerations. Proceeding of 1st Asia Symposium on Quality Electronic Design, pp: 313-318.

Ning, X. and J. Zhonghua, 2008. Thermal aware floorplanning using Gauss-Seidel method. J. Electron., 25(6): 845-851.

Sankaranarayanan, K., S. Velusamy, M. Stan and K. Skadron, 2005. A case for thermal-aware floorplanning at the micro architectural level. J. Instr. Level Parall., 7: 1-16.

Schafer, B.C. and T. Kim, 2008. Hotspots elimination and temperature flattening in VLSI circuits. IEEE T. VLSI Syst., 16(11).

Skadron, K., M. Stan, W. Huang, S. Velusamu, K. Sankaranarayanan and D. Tarjan, 2004. Temperature aware micro architecture: Modeling and Implementation. ACM T. Archit. Code Op., 1: 94-125.

Tsai, C.H. and S.M. Kang, 2000. Cell-level placement for improving substrate thermal distribution. IEEE T. Comput. Aid. D., 19(2): 253-266.