Thermal Analysis and Modeling of 3D Integrated Circuits for Test Scheduling

Indira Rawat
Department of Electrical Engineering, Govt. Engineering College, Ajmer, Rajasthan , India.
indirarawat@gmail.com

M.K. Gupta
Deptt. Of Electronics and Communication Engg., Maulana Azad National Institute of Technology, Bhopal, India.
gupta4@yahoo.com

Virendra Singh
Deptt. Of Electrical Engineering, Indian Institute of Technology,Bombay Mumbai, India
virendra@computer.org

Abstract

A lot of present research is being devoted to the adoption of a new technology called 3D integration. The demand of today is also to enable the heterogeneous integration of different technologies so that a true SoC design can be realised. The technology poses several problems, the prominent ones being the removal of heat. The testing issue also comes into picture. One of the possible solutions to the problems can be the spliting of the inner stacks into more than one group and testing of these first so as to take advantage of low ambient temperature in the beginning of testing. The floorplan has been adjusted and the effect on temperature studied.

1. Introduction

The challenges of testing a 3DIC are enormous considering especially the thermal problem. The need of today is to have smaller, cheaper and more reliable chips. The performance and cost of VLSI is being governed by the interconnects due to decreasing wire pitch and increasing die size. There is also a need of integration of disparate signals and technologies which are offered by SoC design concept. The planar 2D technology is not suitable for that. It is realised that 3D integration is a technology which can significantly improve deep submicron interconnect performance, packing density of transistors is improved and power dissipation is reduced. The 3D design refers to the vertical stacking of the various blocks of a chip. There can be a multiple level of interconnects. The various layers can be connected together by vertical layers of interconnects. The technology can be exploited to build SoCs by placing circuits with different voltages and performance requirements in different layers.

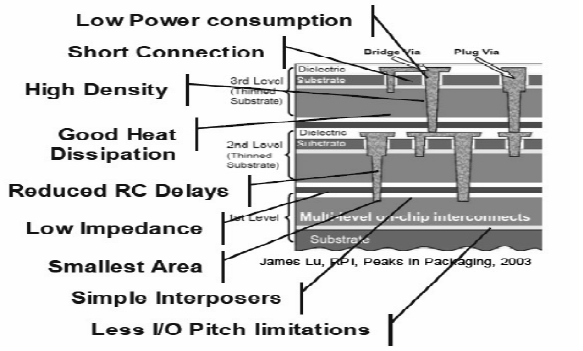

Figure 1. A general 3D structure

Any new technology adoption brings with it various challenges. There are many 3D technology out of which TSV based seems to be the most promising one. The TSV based stacking can be done by two main technologies. (1) Face-to-face (F2F) bonding where two wafers (dies) are stacked with their faces on. (2) Face to back (F2B) where substrate of one is bonded to face off other. The TSVs are the key components of a 3DSIC. They are used for providing test access to logic blocks on different layers. The TSVs too require testing. Thermal and power consideration problem is also one of the predominant ones in 3D. Test access, test scheduling problems need to be addressed. A generalised 3D structure is shown in Figure1[13]. Shukri J. Souri K. Banerjee et.al.[4] have presented an analytical treatment of ICs with multiple Si Layers (3DICs). They showed that a

978-1-4673-6484-3/13/$31.00 ©2013 IEEE

significant improvement in performance about 145% and area is possible by dividing a chip into separate blocks and placing in different stacks. 3D Architecture presents flexibility in design, placement and routing. The concerns of 3D integration have also been discussed in their paper. Thermal Analysis of vertically connected circuits was discussed by Michael B. Kleiner et .al. [2] using a 1-D model. The heating behavior of individual MOSFETs was studied to investigate local heating

2. Thermal Analysis

In this paper an attempt is made for thermal analysis and modeling of 3D stacked ICs. For the background it is worth mentioning that thermal analysis of vertically arranged circuits was done way back in 1960[5] when James Early of Bell laboratories discussed 3D stacking and predicted that the main hurdle to its adoption would be the heat removal. The attraction of such an arrangement included the high packaging density obtainable, ability to integrate different technologies in one device. A 3D integrated circuit is assembled by stacking individual chips which can be connected by different technologies out of which TSV prove to be the most attractive one. As with any new technology there are drawbacks associated with it as well, prominent ones being the thermal issues.

An extremely important issue in 3D ICs is the heat dissipation problem. As 3D IC results in reduction in chip size, if a 2D ckt is transformed into a 3D one obviously with a smaller area, the power density is now more due to same heat production in a smaller area. Device performance can degrade and results in reduction in chip reliability due to increased junction leakage. Electro migration and other failure mechanism can also result. For looking into the heat/thermal problem, we take a look at the analysis available in the literature [1, 4, 10]

2.1 Package Thermal Resistance Model

The relation between die temperature rise ΔT_{die} and P is given by [1]

$$\Delta T_{die} = (T_{die} - T_{amb}) = P.\ R_{\Theta} \qquad (1)$$

Where P is the power dissipation and R_{Θ} is the effective thermal resistance from the Si devices to the heat sink and mostly due to package material.

It can be expressed as

$$R_{\Theta} = \frac{R_n}{A} \qquad (2)$$

$R_n = \left(\frac{tsi}{Ksi}\right) + \left(\frac{tpkg}{Kpkg}\right)$; where Si values are for silicon and pkg values are for package, thickness and conductivity represented by t and K.

$$\Delta T_{die} = P.\ R_n/A \qquad (3)$$

Based on this an idea of the temperature of the die can be had. Value of R_n available from published data is 4.75° C/ W-cm$^{2.}$. The value of power dissipation has been obtained from data by ITRS. Table 1 and Table 2 show the temperature rise when area is changed for two cases for same power dissipation. As Area is reduced there is a sharp rise in die temperature. Approximately for the same power dissipation, increase in temperature is about 25% more.

Table 1-Temperature rise w.r.t. area (case 1)

P(W)	Temperature ° C	Area mm^2
100	118.75	4
80	95	4
75	89	4
70	83.12	4
60	71.25	4

Table 2- Temperature rise w.r.t. area (case 2)

P(W)	Temperature ° C	Area mm^2
100	158	3
80	126	3
75	118.75	3
70	110.83	3
60	95	3

The above data refer to the 2D circuits. The problem becomes all the more acute for 3D circuits.

In [8,12] full –chip temperature calculation has been discussed. The devices in the substrate form the main source of heat generation. The interconnects also get heated up and contribute towards heat generation. Operating temperature of a VLSI chip can be calculated from

$$T_{chip} = T_a + R_{\Theta}\ P/A \qquad (4)$$

Where T_{chip} is the average chip (silicon junction) temperature, T_a is the ambient temperature = 25ºC, P is the total power consumption in W. This is consistent with (1).

2.2 Analytical Die Temperature Model

The temperature rise of the j_{th} active layer in an n-layer 3D chip can be expressed as [1]

$$\Delta T_j = \sum_{i=1}^{j} [R_i (\sum_{k=i}^{n} \frac{P_k}{A})] \quad (5)$$

Where n is total number of active layers;
R_i is thermal resistance between i^{th} and $(i-1)^{th}$ layers;
P_k is power dissipation in k^{th} layer.

Assuming identical power distribution between each layer and identical thermal resistances (R) between layers, the temperature rise of uppermost (n^{th}) layer in an n-layer 3D chip can be expressed as

$$\Delta T_n = \left(\frac{P}{A}\right)\left[\frac{R}{2}n^2 + \left(R_1 - \frac{R}{2}\right)n\right] \quad (6)$$

R_1 is mostly due to package thermal resistance between first layer and heat sink.
The models available so far relates to the temperature of the complete stack or die i.e. they take the temperature of a complete die as same. Practically however this is not the case as different cores dissipate different amount of heat or power. Our work is based on finding the temperature of individual cores which constitute a single stack.

2.3 Analysis based on three dimensional heat conduction

The method of Fourier's law of conduction of heat has been used here. Conduction is the main mode of heat transfer in solids and this is the mode considered in this work. Heat flow by conduction is governed by Fourier law.

$$q = -k\frac{dT}{dx} \quad (7)$$

Where k is the thermal conductivity of the material in (W/m.K). This equation says that heat flux q, (the flow of heat per unit area per unit time) at a point in a medium is directly proportional to temperature gradient at that point. The minus sign indicates that heat flows in the direction of decreasing temperature. If q is written as Q/A where Q is the heat transfer rate, A is the conducting area, and L is the length of material, now equation becomes:

$$Q = -kA\frac{(T_1 - T_2)}{L} \quad (8)$$

As in our previous work [11] it is discussed here for three dimensional cases. The conduction mode of heat transfer is the most prominent one in solids. Here the heat transfer in three dimensions will take place from one core to the next. A three dimensional core will have a maximum of six neighbours. When a core is tested it produces a lot of heat. The heat will be transferred to its neighbours by conduction. Based on the temperature of core the algorithm will determine the next core to be tested. In our previous published work, we had done sequential testing where parallelism was not done. Now we have done concurrent testing of cores.

We consider test scheduling of a 3D stack built by SoC d695as shown in Figure 2 where stacks can be built using different floorplans. The cores on the floorplan will have neighbors on all sides. Each core can have a maximum of six neighbors. We have neglected the effect of thermal interface material between the two stacks. When a core is tested, heat is generated which spreads to all its neighbors both vertically and horizontally [9].

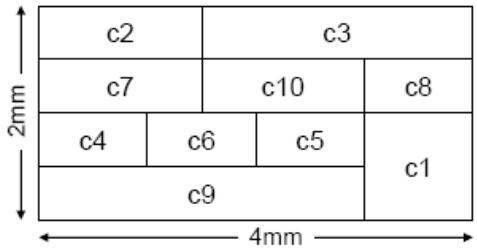

Figure 2. Floorplan of soc d695

For three dimensional case consider a cell (i,j,k) with side length Δx_i, Δy_i ,Δz_k . At most there will be 6 adjacent neighbors for each cell in 3D case. The heat flow $Q_{i,j,k+1/2}$ (W) from one cell (i,j,k) to cell above (i,j,k+1) is given by conductance between the two cells multiplied by the temperature difference between the two cells. Similarly heat flow from cell (i,j,k) to other six cells in its proximity are given by following equations.

$$Q_{i,j,k+\frac{1}{2}} = K_{i,j,k+\frac{1}{2}}(T_{i,j,k} - T_{i,j,k+1})(W)$$
$$Q_{i,j,k-\frac{1}{2}} = K_{i,j,k-\frac{1}{2}}(T_{i,j,k} - T_{i,j,k-1})(W)$$
$$Q_{i-\frac{1}{2},j,k} = K_{i-\frac{1}{2},j,k}(T_{i,j,k} - T_{i-1,j,k})(W)$$
$$Q_{i+\frac{1}{2},j,k} = K_{i+\frac{1}{2},j,k}(T_{i,j,k} - T_{i+1,j,k})(W)$$
$$Q_{i,j+\frac{1}{2},k} = K_{i,j+\frac{1}{2},k}(T_{i,j,k} - T_{i,j+1,k})(W)$$
$$Q_{i,j-\frac{1}{2},k} = K_{i,j-\frac{1}{2},k}(T_{i,j,k} - T_{i,j-1,k})(W)$$

We have built up a 3D-SoC by stacking several layers of SoC d695[7]. The d695 consists of 10 cores. The die size taken is 4mm X 4mm. In the proposed temperature aware test scheduling scheme, the following points are considered.

1. Start initially with the testing of innermost core to have the advantage of low temperature of neighboring cores.
2. For the next step, the cores with the same temperature can be grouped in a common task that is parallel testing is done.
3. Testing produces heat which spreads to other cores.
4. The temperature rise of all cores is calculated. The cores with least temperature rise are selected for testing on the same stack i.e. parallel testing of cores on same stack is carried out. They are deleted from the list of cores to be tested.
5. This way testing scheduling of all cores is done till all cores have been tested.
6. Since temperature also has some dependence on design, floorplanning can be adjusted. The innermost layer is split up into two as shown in Table 3. As observed the temperature rise of various combinations can be outlined as shown in Table 4.

4. Conclusion and future work

Table 3 shows the various combinations considered. The first one consists of 3 layers, the number of cores being 5 in first layer, 5 in second layer and 10 in third layer. The highest temperature difference of scheduled and sequential testing differs by 3K. But in the third combination where two 5core layers are trapped in between layers of 10 cores, the difference is 23K.

The largest core should be stacked next to heat sink. The test scheduling has been done based on temperature. Due to a large amount of heat produced, there has to be some means of heat removal. It has been highlighted in [6] by Muhannad S. Bakir, Calvn King et al. and [3] by B. Goplen et.alThis will form the basis of our future course of work on cooling methods. The stress is to be laid upon the placement of thermal vias.

Table 3- Different floorplan options for 3 and 4 layers each consisting of a d695 and dissipating equal power in testing

Combination	layers	No. of cores
1	1 2 3	5 5 10
	1 2 3 4	5 5 10 10
3	1 2 3 4	10 5 5 10

Table 4 – Thermal performance of different floorplans

	Scheduled		Sequential	
Combination	Highest (K)	Lowest (K)	Highest (K)	Lowest (K)
1	342	334	345	334
2	338	333	342	335
3	361	340	384	345

References

[1] K. Banerjee, S.L.Souri, P. Kapur and K. C . Saraswat, "3D ICs: A novel chip design for improving deep submicrometer interconnect performance and system on chip integration", *Proc. IEEE, 89(5)*, 2001, pp 602-633.

[2] M. B. Kleiner, S.A. Kuhn, P. Ramm, W. Weber, "Thermal Analysis of vertically integrated circuits", *IEDM Tech. Digest,* 1995, pp. 487-490.

[3] B. Goplen and S. Sapatnekar, "Thermal Via Placement in 3D ICs", *Proc. Intl. Symp. on physical design*, 2005, pp 167-174.

[4] S. J. Souri, K. Banerjee, A. Mehrotra and K. C.Saraswat, " Multiple Si layer ICs: Motivation, performance analysis and design implications", *37th ACM design Automation Conference*, 2000, pp 873-880.

[5] J. Early, "Speed, Power and Component Density in Multielement High Speed Logic System", *Proc. IEEE International Solid State Circuits Conf,* Feb. 1960, pp 78-79.

[6] Muhannad S. Bakir, Calvn King et al, "3D Heterogeneous Integrated Systems: Liquid Cooling, Power Delivery and Implementation" *IEEE Custom Integrated Circuits Conf, (CICC)*, Oct. 2008, pp 663-670.

[7] F. A. Hussin, Thomas Edison, et. Al, "Red SOCs 3D: Thermal –safe Test Scheduling for 3D-stacked SOC", *Proc.*

Asia-pacific conference on circuits and systems(APCCAS 2010) pp 264-267.
[8] Kevin Skadron , T. Abdelzaher and Mircea Stan, , "Control Theoretic and Thermal RC Modeling for Accurate and Localised Dynamic Thermal management", *Proc. 8th IEEE Symposium on High Performance Computer Architecture,* 2002, pp 17-28.
[9]. Thomas Blomberg, "Heat conduction in two and three dimensions", *Report TVBH*, 1996.
[10]. W. Huang, S. Ghosh, S. Velusamy, K. Sankaranarayanan, K. Skadron, and M. R. Stan, "Hotspot: A Compact Thermal Modeling Methodology for Early-Stage VLSI Design", *IEEE Transactions on VeryLarge Scale Integration (VLSI) Systems,* 14(5): May 2006, pp 501–513.
[11] I. Rawat, M. K. Gupta and V. Singh, "Scheduling Tests for 3D SoCs with thermal Constraints", *Proc. 10th IEEE EWDTS*, Sept. 2012.
[12] M. Pedram and S. Nazarian, "Thermal Modeling,Analysis and Management in VLSI circuits; Principles and Methods", *Proc. IEEE, vol.94, issue 8*, pp 1487-1501.
[13] James Lu, "Peaks in Packaging ",in *Proc. 3D technical Symposium Nov. 2007,* Paris , France pp3.