TSV Stress Aware Timing Analysis with Applications to 3D-IC Layout Optimization

Jae-Seok Yang, Krit Athikulwongse*, Young-Joon Lee*, Sung Kyu Lim*, and David Z. Pan
Department of ECE, University of Texas at Austin, Austin, TX USA
*School of ECE, Georgia Institute of Technology, Atlanta, Georgia, USA
jsyang@cerc.utexas.edu, {krit, yjlee, limsk}@ece.gatech.edu, dpan@ece.utexas.edu

ABSTRACT

As the geometry shrinking faces severe limitations, 3D wafer stacking with through silicon via (TSV) has gained interest for future SOC integration. Since TSV fill material and silicon have different coefficients of thermal expansion (CTE), TSV causes silicon deformation due to different temperatures at chip manufacturing and operating. The widely used TSV fill material is copper which causes tensile stress on silicon near TSV. In this paper, we propose systematic TSV stress aware timing analysis and show how to optimize layout for better performance. First, we generate a stress contour map with an analytical radial stress model. Then, the tensile stress is converted to hole and electron mobility variations depending on geometric relation between TSVs and transistors. Mobility variation aware cell library and netlist are generated and incorporated in an industrial timing engine for 3D-IC timing analysis. It is interesting to observe that rise and fall time react differently to stress and relative locations with respect to TSVs. Overall, TSV stress induced timing variations can be as much as ± 10% for an individual cell. Thus as an application for layout optimization, we can exploit the stress-induced mobility enhancement to improve timing on critical cells. We show that stress-aware perturbation could reduce cell delay by up to 14.0% and critical path delay by 6.5% in our test case.

Categories and Subject Descriptors

B.7.2 [**Hardware, Integrated Circuit**]: Design Aids

General Terms

Design

Keywords

3DIC, TSV, stress, mobility variation, timing analysis

1. INTRODUCTION

3D-IC stacking has gained tremendous interests for IC integration in order to reduce wire length and increase density [1–4]. TSVs are inserted for wafer-to-wafer connection in 3D-ICs. Tungsten(W), poly-silicon, and copper(Cu) have all been considered as fill materials of TSVs. Since copper has low resistivity, it is widely used material for TSV fill. However, copper CTE differs from silicon CTE which can be a source of silicon strain. CTE of copper is $17\times10^{-6}K^{-1}$ at $20^{\circ}C$, while CTE of silicon is $3\times10^{-6}K^{-1}$ at $20^{\circ}C$ [5]. The CTE mismatch between copper and silicon causes inevitable stress on silicon near TSVs. Because copper electroplating and annealing temperature is higher than operating temperature, tensile stress appears on silicon [6, 7] after cooling down to room temperature.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
DAC'10, June 13-18, 2010, Anaheim, California, USA
Copyright 2010 ACM 978-1-4503-0002-5 /10/06...$10.00

The tensile stress on silicon causes reliability problem such as cracking. In addition, the stress can change mobility of carriers. Therefore, TSV stress induced by CTE mismatch may cause timing violation if cells on a critical path are placed near TSVs. Tensile stress enhances electron mobility. However, hole mobility is either enhanced or degraded depending on TSV and transistor channel direction. Longitudinal tensile stress reduces hole mobility while transverse tensile stress increases the mobility [8]. When TSV induced tensile stress is 100MPa and the stress works for longitudinal direction, hole mobility degradation can be up to 7.2%, which makes PMOS transition slow. If PMOS is on a critical path, it can cause unexpected setup time violation which is not detected with the current timing analysis flow.

Even though several papers have been published regarding TSV stress for reliability, this is the first paper addressing TSV stress from the circuit design perspective, to our best knowledge. In this paper, we propose a design flow to analyze timing variation by TSV induced stress, and show its implications for layout optimizations during 3D-IC design. The first step of our framework is to generate stress map which is used to estimate hole and electron mobility variation. Stress calculation is based on analytical model and linear super-position. Since every cell near TSVs has a different mobility depending on stress and orientation between channel and TSV, we substitute a cell near TSVs to another cell having the same topology but having different timing characteristics according to the estimated hole and electron mobility change.

To show the benefit of our framework, we present that TSV stress aware design plays an important role to optimize timing by adjusting cell locations to take advantage of enhanced mobility property due to TSV stress. Since hole mobility contour differs from electron mobility contour, PMOS and NMOS should be optimized separately. If a PMOS transistor in a cell is on a critical path, the cell becomes a critical cell for hole mobility optimization. An NMOS critical cell can be optimally placed in a similar manner.

The main contributions of this paper include the following:

- We show that TSV stress can change hole mobility quite significantly, e.g., from -22% to +10%, which means more than 20% variation for single cell delay. Thus it can deteriorate the overall chip performance thus must be considered during timing analysis and optimization.
- We propose several layout optimization techniques including small perturbation and optimal cell rotation, and show that the optimization can improve single cell delay by 14% and critical path delay up to 6.5%

The rest of the paper is organized as follows. First, the overall stress-aware timing analysis and design flow is presented in section 2. We propose compact mobility modeling in section 3. In section 4, we will explain how to analyze timing with TSV stress. Experimental results are shown in section 5, followed by conclusion in section 6.

2. TSV STRESS AWARE DESIGN FLOW

Overall flow for 3D-IC design methodology is proposed in Fig. 1. Our timing analysis consists of two steps. The first step is to calculate TSV induced stress and mobility change.

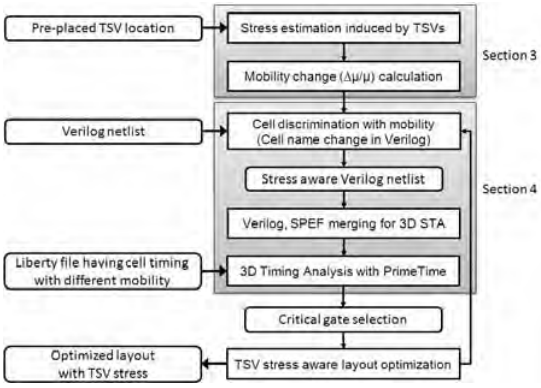


Figure 1: Overall flow for TSV stress aware design.

Since FEA simulation which provides an accurate solution takes several hours to simulate stress for one TSV, we use the analytical model proposed in [6]. Mobility(μ) change as a function of applied stress(σ) has been proposed by the following formula [9].

$$\frac{\Delta\mu}{\mu} = -\Pi \times \sigma \tag{1}$$

where Π is the tensor of piezo-resistive coefficients, and σ is the applied stress in silicon. Positive σ means tensile stress while compressive stress is represented by negative σ. We will explain the stress and mobility modeling in section 3. The second step is 3D timing analysis with TSV stress. We use PrimeTime as a STA (static timing analysis) engine. In section 4, we explain how to deal with verilog netlist and timing library to consider mobility variation. The timing result can be used for layout optimization. Intuitively, if a PMOS in a cell is on a critical path, the cell should be moved to hole mobility enhanced zone. Then, we can run timing analysis iteratively to verify perturbation effects.

3. MOBILITY VARIATION MODELING

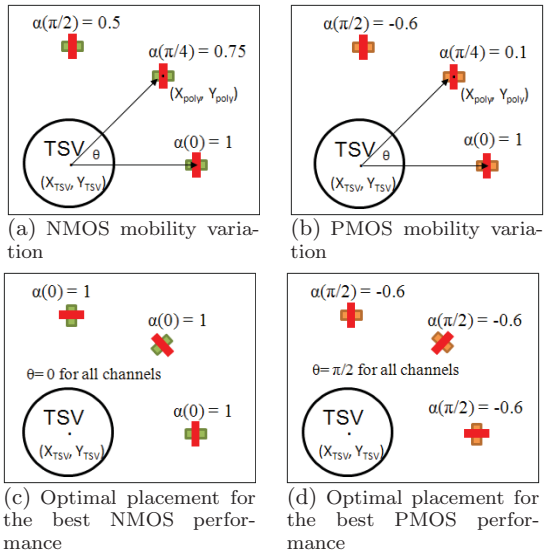
In this paper, we assume that TSV shape is cylindrical, which is widely used for better manufacturability. FEA based TSV simulation has been proposed [5,6]. The simulation approaches provide an accurate solution with long runtime which is not acceptable for our design flow that should calculate stress for several thousands of TSVs iteratively after each optimization. Assuming 2-D radial plain stress, we use the following analytical solution which is known as $Lame'$ stress solution in [6].

$$\sigma_{rr} = -\frac{B\Delta\alpha\Delta T}{2}\left(\frac{R}{r}\right)^2 \tag{2}$$

The analytical stress model provides a relatively accurate solution [6]. In the formula (2), B is biaxial modulus, $\Delta\alpha$ is CTE difference between copper and silicon, ΔT is the temperature difference between copper annealing and operating temperature. R is TSV radius, and r is a distance from the center of TSV. We assume that ΔT is $175^{\circ}C$ which is the case of $25^{\circ}C$ for the room temperature and $200^{\circ}C$ for the copper annealing temperature [10]. The formula shows that the thermal stress near TSV depends on the ratio of TSV radius and a distance from a TSV edge.

The formula (1) provides an efficient way to calculate mobility variation due to σ_{rr}. Mobility change depends on not only σ_{rr} but also orientation between applied force and a transistor channel. The empirical value for showing the relation of mobility change and a channel direction has been proposed in [11]. We extend the formula (1) to consider stress and channel direction in (3).

$$\frac{\Delta\mu}{\mu}(\theta) = -\Pi \times \sigma_{rr} \times \alpha(\theta)$$



(a) NMOS mobility variation

(b) PMOS mobility variation

(c) Optimal placement for the best NMOS performance

(d) Optimal placement for the best PMOS performance

Figure 2: Optimal orientation of MOSFET to maximize mobility for (001) surface, ⟨110⟩ channel.

$$\theta = \tan^{-1}\left|\frac{Y_{TSV} - Y_{poly}}{X_{TSV} - X_{poly}}\right| \tag{3}$$

where $\alpha(\theta)$ is an orientation factor as a function of θ which is defined as degree between the center of TSV and the center of a transistor channel when a transistor is placed vertically as shown in Fig. 2(a),(b). Π is the piezo-resistive coefficient at $\theta = 0$ which works as longitudinal stress.

In Fig. 2(a), if NMOS is in right side of TSV, θ becomes zero, and $\alpha(0)$ becomes one, which enhances NMOS mobility at its maximum. However, if NMOS is in upper side of TSV, $\alpha(\pi/2)$ is 0.5, which means that NMOS mobility increase is half of the enhancement at θ=0. PMOS shows opposite trends, which has the best mobility enhancement at θ=$\pi/2$. Fig. 2(c) and (d) show the transistor direction for the best performance. Even though the mixed channel direction is not allowed due to the patterning difficulty, the observation provides a way to optimize layout for 3D-ICs.

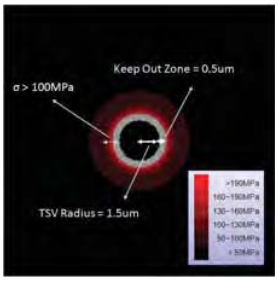
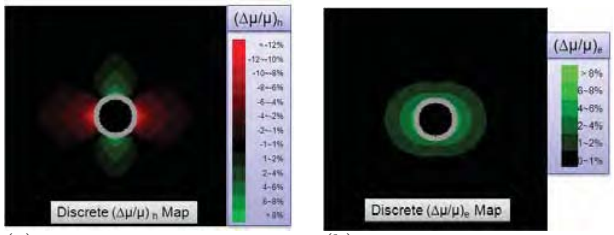


Figure 3: Stress contour map with 0.5um KOZ.

We generate stress contour map based on (2). Fig. 3 shows contour for a TSV having radius of 1.5um. Since the region near TSV may have a crack or extremely high stress, we define that 0.5um from TSV edge is Keep-Out-Zone(KOZ), in which no cell is allowed to be placed. We can see stress of more than 200MPa out of KOZ. Approximately, 100MPa stress appears on the region of 1um from a KOZ edge.

Fig. 4(a) shows a contour map for hole mobility variation. From the contour, we can see that hole mobility decreases in a horizontal direction, while it increases in a vertical region. 45° direction has no hole mobility change. Contour map for electron mobility variation is presented in Fig. 4(b). As we see in Fig. 2(a), horizontal direction has more mobility



(a) Contour map for hole mobility variation (b) Contour map for electron mobility variation

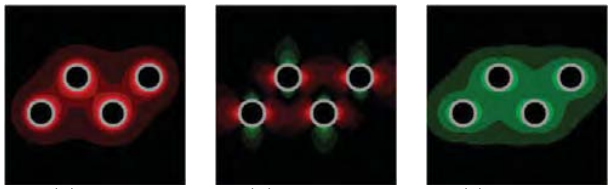
Figure 4: Mobility contour map for a TSV.

enhancement zone.

Since many TSVs for signaling, power/ground and clock network are used for 3D-ICs, our variation model needs to be extended to stress effect for multiple TSVs. Each TSV works as stress source to silicon. When a position in a wafer is strained by multiple stress sources, linear super-position can provide the multiple stress solution [6]. We propose the mobility variation for multiple TSVs.

$$\left.\frac{\Delta\mu}{\mu}\right|_{total} = \sum \frac{\Delta\mu}{\mu}(\theta) = -\Pi \sum_{i \in TSVs} (\sigma_i \times \alpha(\theta_i)) \qquad (4)$$

where σ_i is the tensile stress caused by i^{th} TSV, $\alpha(\theta_i)$ is the orientation factor of i^{th} TSV. θ_i is the degree between the center of i^{th} TSV and a point that we want to get mobility variation.



(a) σ_{total} (b) $(\Delta\mu/\mu)_h$ (c) $(\Delta\mu/\mu)_e$

Figure 5: Linear super-position of TSV stress.

Fig. 5 shows stress and mobility variation contour with linear super-position for four-TSV array. We can see more stress in a region between TSVs. $(\Delta\mu/\mu)_e$ contour has similar trend with stress contour. However, $(\Delta\mu/\mu)_h$ has less variation between TSVs because zigzag TSV placement has compensation effect for positive and negative hole mobility between adjacent rows.

4. STRESS AWARE TIMING ANALYSIS

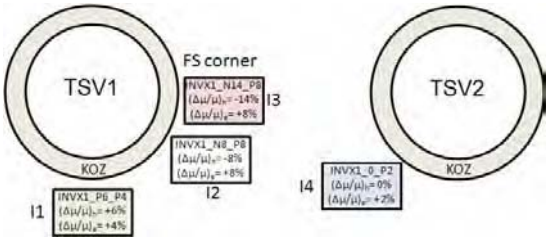


Figure 6: Timing corner determination according to mobility variation.

Even though topology of a cell is the same, its timing characteristic will be different depending on stress amount and stress direction to transistor channel. Fig. 6 shows the example that cells having the same topology and size are in different timing corners systematically determined by TSVs. When two TSVs are near three inverters, cell characteristics are different in a different position. From the formula (4), we can determine $\Delta\mu/\mu$ in any point for a given layout. After mobility calculation, our framework renames for cells to include mobility variation in verilog netlist. For example, I2 is renamed to INVX1_N8_P8 which means -8% hole mobility, +8% electron mobility in Fig. 6.

We prepare a verilog netlist and a parasitic extraction file (SPEF) for each die. In addition, we make a top level verilog netlist that instantiates the dies and connects them using wires which corresponds to TSV connections. Then we make a top level SPEF file for the TSV connections. With a proper timing constraints file, we can run PrimeTime and get the 3D STA results.

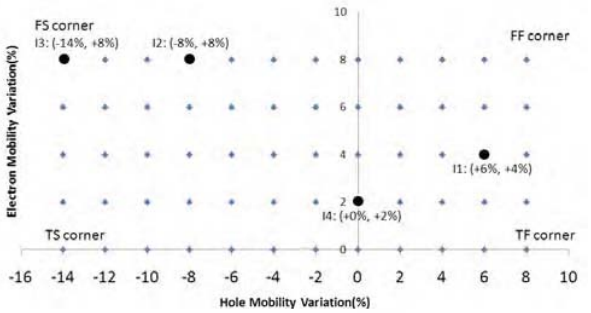
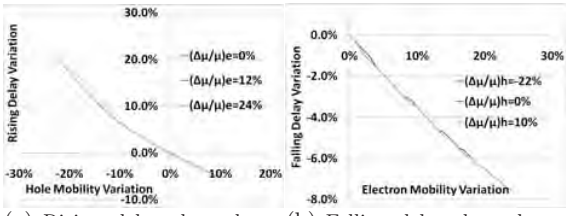


Figure 7: Timing corner with TSV stress.

To consider the systematic variation during timing analysis, we characterize a cell with different mobility corners as shown in Fig. 7. Hole mobility variation is from -14% to +8%, and electron mobility variation is up to +8% to cover stress caused by TSVs in Fig. 6. I1 in Fig. 6 is matched the corner near FF corner, while I3 is in FS corner. With mobility variation aware library and verilog netlist having renamed cells, we run PrimeTime to do timing analysis with TSV stress.



(a) Rising delay dependency on hole mobility variation (b) Falling delay dependency on electron mobility variation

Figure 8: Inverter variation with mobility change.

To cover mobility variation caused by multiple TSVs, the mobility variation range needs to be extended to PMOS: from -22% to +10%, NMOS: from 0% to +24%. If mobility step is 2%, we need to characterize 221 library with different mobility values which is not available. However, we can observe that rising delay variation only depends on $(\Delta\mu/\mu)_h$, falling delay variation depends on $(\Delta\mu/\mu)_e$ from Fig. 8. When we simulate inverter rising delay with mobility variation, electron mobility variation does not work for the delay. Similarly, we can see that falling delay only depends on electron mobility variation. In addition, from Fig. 8, we can see that hole mobility variation can cause more than 20% PMOS performance variation depending on device technology, and electron mobility variation can enhance NMOS performance up to 7.5%. We use inverter in NCSU library and PTM spice model [12] to obtain Fig. 8. Therefore, we can fix $(\Delta\mu/\mu)_e$ when we sweep $(\Delta\mu/\mu)_h$. 30 (=17+13) library characterization will be enough to cover the entire mobility set. If mobility step is 4%, 16 (=9+7) library set is required. Since delay variation has semi-linear dependency on mobility variation, we can use interpolation for the mobility value between two libraries.

5. EXPERIMENTAL RESULTS

We implement TSV stress aware 3-D timing analysis flow in C++ and generate the mobility aware library based on NCSU 45nm cell library. TSV used in this experimentation is in Table 1.

Table 1: TSV specification

Width	Landing pad	KOZ	Height	Dielectric	Resistance	Capacitance
4.14um	4.54um	0.4um	20um	0.2um	0.1Ω	70fF

Table 2: Longest path delay and TNS comparison

Circuit	#Cells	Without TSV stress		With TSV stress		Difference	
		Longest Delay(ns)	TNS (ns)	Longest Delay(ns)	TNS (ns)	Longest Delay	TNS (ns)
IDCT	14,864	12.07	-21,293	11.91	-19,652	-1.32%	-7.71%
8051	15,712	4.78	-7,868	4.94	-7,956	3.32%	1.12%
8086	19,895	9.56	-8,557	9.56	-9,045	0.00%	5.71%
MAC2	29,706	7.72	-17,561	7.72	-17,619	0.07%	0.33%
ETHERNET	77,234	18.30	-476	18.95	-482	3.58%	1.24%
RISC	88,401	8.28	-1,249	8.34	-1,535	0.74%	22.90%
B18	103,711	11.28	-2,082	11.25	-1,823	-0.27%	-12.43%
DES_PERT	109,181	8.61	-2,801	8.64	-2,575	0.25%	-8.06%
VGA_LCD	126,379	8.01	-543	8.14	-538	1.56%	-1.02%
B19	168,943	13.01	-5,539	12.98	-4,974	-0.20%	-10.20%

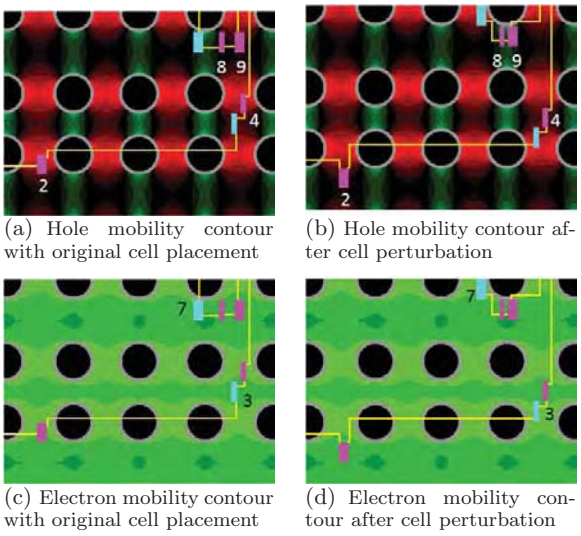
First, we show the efficiency of our compact stress and mobility modeling. Even though we generate mobility contour for a block having die size: $1.75^2 mm^2$, #TSVs: 462, it takes only 14.9s. The proposed timing analysis with compact process/device model is fast enough to be used for iterative optimization purpose.

Second, we compare stress aware timing result with no stress case. Ten benchmark circuits are used to show the timing variation in Table 2. The benchmark circuits are placed for wire length minimization [1] without TSV stress consideration. We assume that there are four dies stacking, and the number of inserted TSVs are 10% of #cells in each circuit. When we consider TSV stress effect, the longest path delay of the benchmarks has variation from -1.32% to 3.58%. Some benchmarks have timing gain while some benchmarks have timing penalty. If we consider TSV stress effect during cells and TSVs placement, we can expect performance improvement for every benchmark. TNS (total negative slack) has more variation from -12.43% to 22.9% which is bigger than delay variation. That motivates the need of TSV stress aware layout optimization.

Last, we manually optimize a critical path in 8051 to present the potential benefit of TSV stress aware layout optimization. Before optimization, the path delay is 4.94ns with stress aware timing analysis. However, we could reduce the delay to 4.62ns with small layout perturbation which is 6.5% improvement. Table 3 shows the gates on the path. We can see the cell renaming according to the mobility variation. We adjust each cell location with small perturbation so that each cell has timing gain. The maximum timing gain in a cell is 14%. Fig. 9 shows how cell relocation works for timing optimization. We capture the placement result on die2 with mobility variation contours. The cells in logic depth 2,4,8 and 9 are hole mobility critical cells because the timing arc is rising on the path. Therefore, we perturb the cells to be placed close to green area in hole mobility contour. However, the cells in logic depth 3 and 7 are electron mobility critical. Therefore, we push the cells to electron mobility enhancement zone in Fig. 9(c) (d).

Table 3: Path optimization with cell perturbation

Logic Depth	Original Gate	Optimized Gate	Timing Arc	Original Delay(ns)	Optimized Delay(ns)	Reduction Ratio
	DFFPOSX1	DFFPOSX1	fall	0.337	0.334	-0.6%
1	NOR3X1_N2_P14	NOR3X1_P4_P14	rise	0.800	0.767	-4.1%
2	AND2X1_N12_P12	AND2X1_P0_P12	rise	0.539	0.492	-8.7%
3	INVX1_N6_P12	INVX1_N6_P16	fall	0.207	0.191	-7.9%
4	INVX1_N12_P12	INVX1_P2_P12	rise	0.653	0.585	-10.4%
5	AND2X1_N16_P16	AND2X1_N4_P14	rise	0.576	0.535	-7.2%
6	BUFX2_P6_P12	BUFX2_P6_P12	rise	0.245	0.216	-11.8%
7	AOI22X1_P4_P10	AOI22X1_P4_P14	fall	0.159	0.148	-7.3%
8	INVX1_P0_P10	INVX1_P2_P12	rise	0.107	0.105	-1.5%
9	OR2X1_N4_P10	OR2X1_P2_P8	rise	0.490	0.468	-4.3%
10	OR2X2_N16_P18	OR2X2_N2_P12	rise	0.068	0.059	-13.3%
11	NOR3X1_P0_P14	NOR3X1_P0_P16	fall	0.100	0.089	-11.6%
12	NAND3X1_N4_P14	NAND3X1_P2_P12	rise	0.055	0.051	-7.3%
13	BUFX2_N4_P14	BUFX2_P4_P12	rise	0.157	0.149	-4.8%
14	OR2X2_P0_P8	OR2X2_P2_P8	rise	0.170	0.169	-1.0%
15	AOI22X1_N16_P16	AOI22X1_N16_P16	fall	0.076	0.075	-1.7%
16	OAI21X1_N4_P14	OAI21X1_P4_P12	rise	0.072	0.069	-4.9%
17	NOR3X1_P2_P14	NOR3X1_P2_P16	fall	0.035	0.034	-2.3%
18	AOI21X1_N18_P18	AOI21X1_P4_P12	rise	0.047	0.040	-14.0%
19	INVX1_N16_P16	INVX1_N16_P18	fall	0.027	0.024	-9.6%
20	OAI21X1_P6_P14	OAI21X1_P6_P14	rise	0.017	0.017	1.6%
		Path Delay		4.93719	4.61777	-6.5%



(a) Hole mobility contour with original cell placement

(b) Hole mobility contour after cell perturbation

(c) Electron mobility contour with original cell placement

(d) Electron mobility contour after cell perturbation

Figure 9: Cell perturbation to take advantage of mobility variation.

6. CONCLUSIONS

The 3D IC stacking requires TSV for interconnection between wafers. Cu TSV causes thermal stress which can lead to significant timing variations. Stress, though commonly believed to have negative impact on timing can actually be taken advantage of for timing optimization, since it is a strongly layout dependent, systematic effect. In this paper, we develop the first-order compact model for mobility variation and propose a design methodology to analyze the systematic variation and optimize layout by locating critical cells in a mobility enhanced region. Our TSV stress-aware timing analysis framework for 3D-IC also opens the opportunity for stress-aware layout optimizations, such as placement and TSV optimizations.

ACKNOWLEDGMENTS

This work is supported in part by NSF and SRC Interconnect Focus Center (IFC).

7. REFERENCES

[1] D. H. Kim, K. Athikulwongse, and S. K. Lim. A Study of Through-Silicon-Via Impact on the 3-D Stacked IC Layout. In *Proc. Int. Conf. on Computer Aided Design*, Nov 2009.

[2] C. Chiang and S. Sinha. The road to 3d eda tool readiness. In *Proc. Asia and South Pacific Design Automation Conf.*, Jan 2009.

[3] Y.-J. Lee, R. Goel, and S. K. Lim. Multi-functional Interconnect Co-optimization for Fast and Reliable 3D Stacked ICs. In *Proc. Int. Conf. on Computer Aided Design*, Nov 2009.

[4] X. Zhao, D. Lewis, H.-H. S. Lee, and S. K. Lim. Pre-bond Testable Low-Power Clock Tree Design for 3D Stacked ICs. In *Proc. Int. Conf. on Computer Aided Design*, Nov 2009.

[5] T. Dao, D. H. Triyoso, M. Petras, and M. Canonico. Through Silicon Via Stress Characterization. In *IEEE International Conference on IC Design and Technology*, 2009.

[6] K. H. Lu, X. Zhang, S.-K. Ryu, J. Im, R. Huang, and P. S. Ho. Thermo-Mechanical Reliability of 3-D ICs containing Through Silicon Vias. In *Electronic Components and Technology Conference*, 2009.

[7] C. S. Selvanayagam, J. H. Lau, X. Zhang, S.K.W. Seah, K. Vaidyanathan, and T. C. Chai. Nonlinear Thermal Stress/Strain Analysis of Copper Filled TSV and their Flip-Chip Microbumps. In *Electronic Components and Technology Conference*, 2008.

[8] S. E. Thompson, M. Armstrong, and C. Auth et al. A 90 nm logic technology featuring strained-silicon. In *IEEE Trans. on Electron Devices*, volume 51, pages 1790–1797, Nov 2004.

[9] C. S. Smith. Piezoresistance effect in germanium and silicon. In *Physical Review*, volume 94, pages 42–49, Apr 1954.

[10] N. Serin, T. Serin, S. Horzum, and Y. Celik. Annealing effects on the properties of copper oxide thin films prepared by chemical deposition. In *Electronic Journals*, volume 20, pages 398–401, May 2005.

[11] H. Irie, K. Kita, K. Kyuno, and A. Toriumi. In-Plane Mobility Anisotropy and Universality Under Uni-axial Strains in n- and p-MOS Inversion Layers on (100), (110), and (111) Si. In *IEEE International Electron Devices Meeting*, 2004.

[12] W. Zhao and Y. Cao. New generation of Predictive Technology Model for sub-45nm early design exploration. In *IEEE Trans. on Electron Devices*, 2006.