Impact of Through-Silicon-Via Scaling on the Wirelength Distribution of Current and Future 3D ICs

Dae Hyun Kim and Sung Kyu Lim
School of Electrical and Computer Engineering
Georgia Institute of Technology
Email: daehyun@gatech.edu

***Abstract*— In this paper, we investigate the impact of TSV scaling on the wirelength distribution of the 3D ICs. This investigation includes wirelength distribution prediction of 3D ICs for current/future process/TSV technologies, studies on the impact of the design granularity at each process node, the impact of the die count, and the impact of TSV area constraint, and cross-comparison among various 2D and 3D technologies.**

I. Introduction

It is expected that 3D ICs outperform 2D ICs mainly due to their reduced wirelength. Recent study showed that 3D implementation of a circuit is superior to its 2D counterpart [1]. One of the basic assumptions for the comparison between 2D and 3D ICs in the literature till now is that the same process technology is used for the 2D and 3D ICs. However, there are several device/interconnect process technologies and TSV technologies available today. Therefore, it is worthwhile comparing 2D and 3D ICs under different process technologies and TSV technologies. For instance, suppose that 3D ICs for a given design are built with $0.18\mu m$ process technology and $3.0\mu m$ TSV, where the TSV size refers to the TSV diameter plus its keep-out distance. Assume that 2D ICs for the same design are built with $90nm$ process technology. In this case, 2D ICs could show better design characteristics such as shorter total wirelength, higher operation frequency, less power, lower cost, and so on. On the other hand, 3D ICs built with $0.18\mu m$ process technology and $1.0\mu m$ TSV for the same design could have better design characteristics than the 2D ICs mentioned earlier.

In this paper, we investigate the wirelength characteristics of various combinations of process technologies and TSV technologies using our TSV-aware interconnect prediction models. According to our simulation results, we draw three important conclusions. First, there exists physically infeasible design space, where the TSV size is not sufficiently small compared to the process technology. Second, the optimal design granularity is strongly dependent on the design parameters such as process technologies, TSV size, die count, area constraint, and so on. Third, 2D ICs built with newer process technologies could be better than 3D ICs built with older process technologies if TSV size is not sufficiently small.

II. Preliminaries

In this section, we discuss the physical feasibility of 3D ICs in terms of the number of TSVs, and review the TSV-aware wirelength prediction model used in this paper.

A. Physical Feasibility of 3D ICs

It has been assumed in the literature that any number of TSVs can be used in 3D ICs. However, there exist lower and upper bounds on the number of TSVs in reality, and ignoring these bounds leads to incorrect analyses. The relation between these two bounds determines the physical feasibility of 3D ICs in terms of the TSV count as follows.

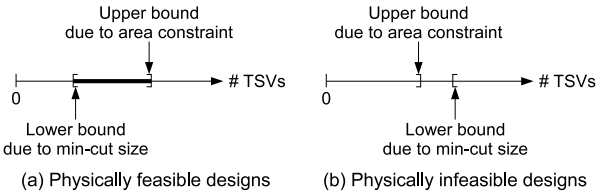

Fig. 1. Relationship between the lower and the upper bound of the TSV count for physically feasible (or infeasible) designs.

The lower bound of the TSV count is determined by the minimum cut size which can be estimated by multi-way min-cut partitioners. If we obtain the minimum cut size for a given design, it is possible to compute the minimum number of TSVs to use from the minimum cut size. For example, suppose that we build a 3D IC using four dies with 10% area unbalancing factor [1]. Then, we can obtain the minimum cut size by running four-way partitioning on the given design. This minimum cut size can be used as a loose lower bound of the number of TSVs.

The upper bound of the TSV count, on the other hand, is determined by the area constraint because inserting more TSVs needs bigger layout area. For example, suppose that total silicon area needed for transistors is A and additional area allowed for TSV insertion is 10% of A. Then upper bound of the TSV count is $0.1A/T$ where T is the area needed for a TSV.

If we obtain the lower and the upper bounds of the TSV count, we can determine whether a given design is a physically feasible design or not under given constraints. Figure 1 shows two examples. In Figure 1(a), the lower bound is smaller than the upper bound and the number of available TSVs is between these two bounds. Therefore this design is a physically feasible design in terms of the TSV count. On the other hand, the lower bound in Figure 1(b) is bigger than the upper bound so there is no available TSV. Therefore this design is a physically infeasible design.

If a design lies in a physically infeasible domain under given constraints, it is possible to make it physically feasible by alleviating area constraint, by using smaller TSVs, or by reducing the die count because using more dies to build a 3D IC usually increases the minimum cut size. However, all of these ways could increase chip cost or decrease chip performance.

This material is based upon work supported by the National Science Foundation under Grant No. CCF-0917000, the SRC Interconnect Focus Center (IFC), and Intel Corporation.

[1]The maximum difference between the maximum and the minimum area of the dies is less than or equal to the unbalancing factor.

978-1-4577-0502-1/11/$26.00 ©2011 IEEE

TABLE I

VARIABLES (CONSTANTS) USED IN OUR EXPERIMENTS

Variable (Constant)	Values
Process technology (nm)	90, 65, 45, 32, 22, 16, 11
Circuit size (# gates)	$10M$
TSV size (μm)	0.1, 0.5, 1.0, $\cdots$, 5.0
# dies	1(2D), 2, 3, $\cdots$, 8
Rent's parameters	α:0.75, k:4.0, p:0.75
Lower bound of # TSVs	$(0.2 \times (n-1))\%$ of # nets for n-die 3D ICs
Area constraint (%)	10 (default), 20, 30
Design granularity ([2])	10 (coarse), 10^4 (fine), and gate-level

B. TSV-aware Wirelength Prediction Model

TSV-aware 3D wirelength prediction model is presented in [2] where its authors improved the non-TSV-aware 3D wirelength prediction model in [3] in order to consider area occupied by TSVs. The primary modification is that a gate pair is not counted during counting the number of gate pairs separated by a distance l if at least one of the two gates in the gate pair lies in TSV spots. The reason of excluding these gate pairs is because it is not possible to place gates in TSV spots so those gate pairs are not realizable.

In addition to consider the TSV area, the model also includes design granularity useful for wirelength prediction of block-level 3D IC design where we design 3D ICs by placing macro modules instead of individual gates. Important results of the paper are 1) non-negligible TSV size affects area and wirelength of 3D ICs significantly, 2) stacking more dies helps reduce wirelength, 3) wirelength of block-level 3D ICs could be shorter than that of gate-level 3D ICs. However, the studies in [2] ignore the physical feasibility discussed in the previous section. In this paper, we consider it and study the relation between gate size and TSV size.

III. CORRELATION BETWEEN PROCESS SCALING AND TSV SCALING

In this section, we present simulation results on correlation of process scaling and TSV sizes. Since there exist many simulation points, we show our simulation settings and methodologies in the first subsection, and proceed to the analysis of simulation results.

A. Simulation Settings and Methodologies

Table I shows variables, constants, and their values used in our simulation.

- **Gate size**: We obtain the average gate size for each process technology from various sources such as standard cell libraries and ITRS.
- **TSV size**: TSV size varies from $5.0\mu m$ which is one of current TSV sizes to $0.1\mu m$ which is a future TSV size.
- **Lower bound of # TSVs**: We have performed multi-way min-cut partitioning on various circuits and obtained loose criteria on the minimum number of TSVs as shown in Table I. For instance, if the number of nets is $2 \cdot 10^6$ and four dies are used, the lower bound of the TSV count is $0.2 \times 3 \times 0.01 \times (2 \cdot 10^6) = 12,000$.
- **Area constraint**: By default, our area constraint is 10%.
- **Design granularity (B in [2])**: By default, we use the finest granularity which means gate-level design. When we vary design granularity, we choose 10 for coarse block-level design and 10^4 for fine block-level design.

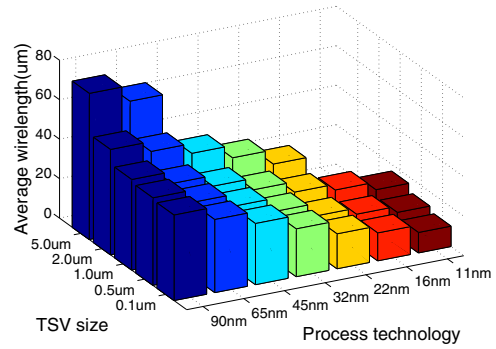

Fig. 2. Wirelength for various process/TSV technologies. # dies: 4.

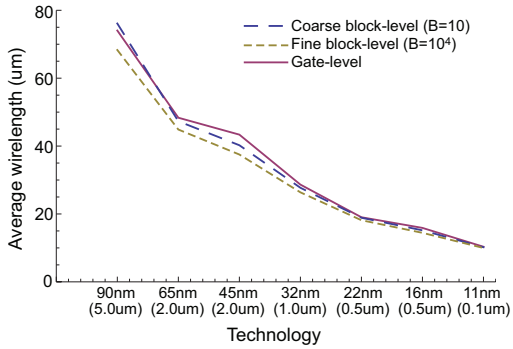

Fig. 3. Impact of design granularity. Numbers in parentheses denote TSV size. # dies: 4.

B. Wirelength Comparison

Figure 2 shows wirelength distributions for various combinations of current/future process and TSV technologies. A general trend is that average wirelength decreases as TSV and device sizes go down. When the TSV size decreases, however, the amount of wirelength reduction also decreases. The reason is that sufficiently small TSV does not cause serious area overhead so that it does not affect wirelength too much.

One thing to notice is the existence of physically infeasible design space (PIDS). For example, combining $5.0\mu m$ TSV with process technologies smaller than $65nm$ creates PIDS when four dies are used as shown in Figure 2 where wirelength is not shown for those combinations. Numerically speaking, for example, the upper bound of the TSV count due to 10% area constraint is approximately $125K$ when the TSV size is $5.0\mu m$ and the process technology is $45nm$. On the other hand, the lower bound of the TSV count is approximately $177K$. Therefore this combination creates PIDS.

When the die count increases from four to eight, more TSVs are needed so there exist more PIDS. Therefore, developing smaller TSVs such as $1.0\mu m$ TSV is necessary for the state-of-the-art and future process technologies to benefit from 3D ICs under tight area constraint.

C. Impact of Design Granularity

Figure 3 shows impact of design granularity for each combination of process and TSV technologies. In the figure, the fine block-level design granularity produces the best wirelength for all technology combinations. On the other hand, the coarse block-level design granularity produces the best wirelength when the die count is

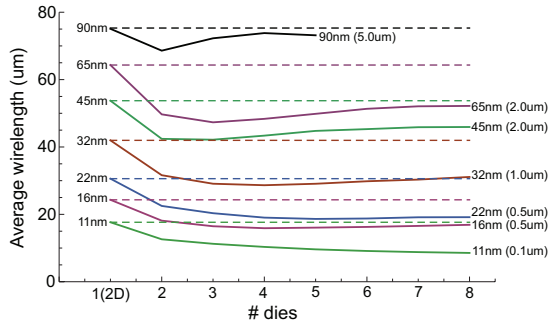

Fig. 4. Impact of die count. Dashed lines are wirelengths of 2D ICs.

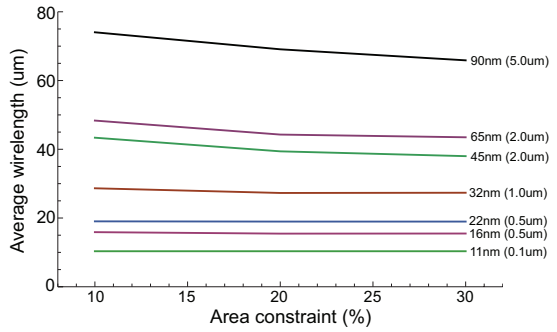

Fig. 5. Impact of TSV area constraint. # dies: 4.

eight (not shown due to page limit). Therefore, we observe that the best design granularity depends on the combination of process and TSV technologies as well as all the design parameters such as the die count and area constraint.

D. Impact of Die Count

Figure 4 shows wirelength of 2D and 3D ICs when the die count varies from one (2D) to eight. Increasing the die count decreases the footprint area, so in general it is expected that wirelength decreases as we stack more dies. However, it depends on the relative sizes between devices and TSVs as shown in the figure. For example, wirelength monotonically decreases as the die count increases when $11nm$ process technology is combined with $0.1\mu m$ TSV under 10% area constraint in the figure. On the other hand, wirelength decreases initially but begins to increase as we stack more dies when $45nm$ process technology is combined with $2.0\mu m$ TSV. Therefore, stacking more dies does not necessarily result in wirelength reduction unless the TSV size is sufficiently small.

Regarding PIDS, design points where the process technology is $90nm$, TSV size is $5.0\mu m$, and the die count is greater than five lie in PIDS as shown in Figure 4.

E. Impact of TSV Area Constraint

Figure 5 shows the impact of TSV area constraint. When the TSV size is big compared to the process technology (e.g., $90nm$ with $5.0\mu m$ TSV), alleviating the area constraint helps reduce wirelength by allowing more TSVs. If the TSV size is very small compared to the process technology (e.g., $22nm$ with $0.5\mu m$ TSV), however, alleviating the area constraint does not result in wirelength reduction because sufficiently many TSVs are already used even under tight area constraint (e.g., 10%).

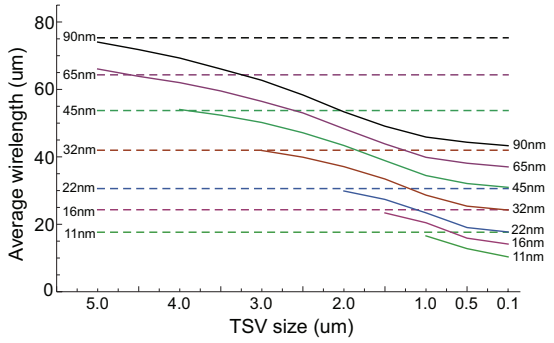

Fig. 6. Cross comparison among various 2D and 3D technologies. Dashed lines are wirelengths of 2D ICs. # dies: 4.

F. Cross-comparison among Various 2D/3D Technologies

We compare various combinations of process and TSV technologies in Figure 6 where we find two interesting results.

First, combining relatively big TSVs (e.g., $4.0\mu m$) with the state-of-the-art or future technologies (e.g., beyond $45nm$) creates PIDS as shown in the figure. Therefore, the state-of-the-art and future process technologies should be accompanied by very small TSV size in order to enable feasible 3D IC designs.

Second, 3D ICs built with an older process technology and a small TSV technology (e.g., $65nm$ with $1.0\mu m$ TSV) could be even better than 2D ICs built with a newer process technology (e.g., $32nm$) as shown in the figure. Similarly, if the TSV size is very small, stacking many dies fabricated with an older process technology (e.g., eight dies and $90nm$ technology) could produce shorter wirelength than 2D ICs built with the state-of-the-art process technology (e.g., $32nm$).

However, this phenomenon can be viewed from opposite aspects. In other words, 2D ICs built with a newer process technology (e.g., $32nm$) could be better than 3D ICs built with an older process technology and a relatively big TSV technology (e.g., $45nm$ with $2.5\mu m$ TSV). Therefore, developing smaller TSVs is extremely important for current and future 3D ICs.

IV. Conclusions

Since TSVs occupy silicon area, there exists the upper bound of the TSV count due to area constraint. Likewise, the minimum cut-size of a design determines the lower bound of the TSV count. These two bounds determine physical feasibility of a 3D IC design for given design parameters and constraints. With this notion and TSV-aware wirelength prediction models, we have demonstrated wirelength distribution of 3D ICs for current/future process/TSV technologies, impact of the design granularity, die count, and TSV area constraint, cross-comparison among various 2D and 3D technologies, and the existence of physically infeasible design space.

References

[1] T. Thorolfsson, K. Gonsalves, and P. D. Franzon, "Design Automation for a 3DIC FFT Processor for Synthetic Aperture Radar: A Case Study," in *Proc. ACM Design Automation Conf.*, July 2009, pp. 51–56.

[2] D. H. Kim, S. Mukhopadhyay, and S. K. Lim, "Through-Silicon-Via Aware Interconnect Prediction and Optimization for 3D Stacked ICs," in *Proc. ACM/IEEE Int. Workshop on System Level Interconnect Prediction*, July 2009, pp. 85–92.

[3] J. W. Joyner, P. Zarkesh-Ha, J. A. Davis, and J. D. Meindl, "A Three-Dimensional Stochastic Wire-Length Distribution for Variable Separation of Strata," in *Proc. IEEE Int. Interconnect Technology Conference*, June 2000, pp. 126–128.