Design Quality Tradeoff Studies for 3D ICs Built with Nano-scale TSVs and Devices

Kaiyuan Yang[1], Dae Hyun Kim[2], and Sung Kyu Lim[2]
[1]Institute of Microelectronics, Tsinghua University, Beijing, China
[2]School of Electrical and Computer Engineering, Georgia Institute of Technology, GA, USA
Email: kyyang6018@gmail.com, daehyun@gatech.edu, limsk@ece.gatech.edu

***Abstract*—Three dimensional integrated circuits (3D ICs) built with through-silicon vias (TSVs) have smaller footprint area, shorter wirelength, and better performance than 2D ICs. However, the quality of 3D ICs is strongly dependent on TSV dimensions and parasitics. Using large TSVs may cause silicon area overhead and reduce the amount of wirelength reduction in 3D ICs. In addition, non-negligible TSV parasitic capacitance can result in delay overhead affecting the delay of 3D ICs. Meanwhile, with the development of TSV manufacturing technology, nano-scale TSVs are emerging, which is expected to reduce the overheads caused by using large TSVs. Therefore, this paper investigates the impact of nano-scale TSVs on the quality of 3D ICs at future technology nodes. For this study, we develop a $16nm$ standard cell library, design 3D ICs using different process technologies ($45nm$, $22nm$, and $16nm$) and various TSVs diameters (from $5\mu m$ to $0.1\mu m$), and discuss the impact of nano-scale TSVs.**

I. Introduction

Three dimensional integrated circuits (3D ICs) using through-silicon vias (TSVs) have emerged as a promising technology to enable low-power, high-performance integrated circuits. By vertically stacking multiple dies fabricated separately, 3D ICs reduce the form factor, improve performance, and provide heterogeneous integration. However, the quality of 3D ICs is strongly dependent on the TSV size (diameter) and the TSV capacitance. Recent studies [1]–[6] investigated the impact of TSVs on the quality of 3D ICs in terms of area, wirelength, critical path delay, and power.

According to the recent publications on TSV fabrication technologies [7], [8] and ITRS prediction [9], the TSV diameter will become smaller than $1\mu m$ within a few years. Meanwhile, the state-of-the-art CMOS process technology has reached $32nm$ [10], and more advanced process technologies (e.g., $22nm$) are on their way [11]. Therefore, it is expected that these upcoming processes and nano-scale TSV fabrication technologies will be combined to build future 3D ICs, and several works predicted the quality of future 3D ICs [12], [13].

In [13], a $22nm$ process library is developed to investigate the impact of TSVs on the quality of current and future 3D ICs. However, a $22nm$ process node will be announced in the industry soon and studies with more advanced process technologies beyond $22nm$ are highly demanded. In this paper, we extend the design and analysis work performed in [13] to include $16nm$ process technology. Our contributions in this paper are as follows:

- To generate 2D and 3D IC layouts at $16nm$ process node, we develop a fully functional $16nm$ process and standard cell library based on the prediction of interconnect layers at $16nm$ process node.
- We generate and compare 2D and 3D IC designs with our $16nm$ library as well as $22nm$ [13] and $45nm$ [14] libraries. With these realistic layouts, we study and compare the impact of TSV size and capacitance on current and future 3D ICs.

This material is based upon work supported by the Semiconductor Research Corporation (SRC) under the Integrated Circuit & Systems Sciences (ICSS, Task ID: 2193.001 & 1836.075) and the Interconnect Focus Center (IFC, Theme ID: 2050.001).

The rest of this paper is organized as follows. Section II demonstrates the development flow of our $16nm$ library and shows comparison of three different process libraries. In Section III, we briefly introduce the full-chip 3D IC design and analysis methodologies used in this paper and explain the experimental settings. In Section IV, various experimental results are presented and analyzed. We conclude in Section V.

II. Our $16nm$ Process Library

In this section, we present interconnect layers and standard cells of our $16nm$ library. To develop the $16nm$ standard cell library, we follow the same library design steps described in [13]. We also compare wirelength, area, critical path delay, and power of 2D ICs designed with a $45nm$, a $22nm$, and our $16nm$ technology libraries to show the feasibility of our library.

A. Prediction of Interconnect Layers

We predict the dimensions of the interconnect layers in our $16nm$ library based on the ITRS interconnect prediction [9] and other process technology papers [10], [11], [15], [16]. For example, the metal 1 pitch in the ITRS prediction for $16.8nm$ process is $48nm$ and that for $18nm$ SRAM in [11] is $50nm$. Therefore, we predict that the metal 1 pitch of a $16nm$ process is $46nm$. Table I shows the pitches of eight metal layers in Intel process technologies, a predicted $22nm$ library, and our $16nm$ library. We observe that there exists an almost linear downscaling trend in metal 1 width along with the gate length. Since the aspect ratio of the ITRS prediction for $16.8nm$ process is 1.9, we use the aspect ratio for our $16nm$ library. Table II shows the dimensions (width, pitch, and thickness) of all the metal layers defined in our $16nm$ library. Besides, we apply low-k inter-layer insulator, so the dielectric constant of all the inter-layer insulator layers is set to 1.9, and the dielectric constant of the barrier layers is set to 3.8.

B. Standard Cell Library

After predicting the interconnect layers, we create technology files, which are used to generate standard cell layouts. We draw about 100 cells similar to the $22nm$ library presented in [13]. The placement site width is $0.06\mu m$ and the standard cell height is $0.6\mu m$. The width of the $1\times$ inverter, which is the smallest functional logic gate, is $0.18\mu m$. To visualize the relative size of $45nm$, $22nm$, and $16nm$ standard cells, we show the smallest ($1\times$) two-input NAND gates in Figure 1. After drawing the standard cell layouts, we perform DRC and LVS of each layout. We also extract parasitic RC of each standard cell and characterize the cell to create timing and power

TABLE I
INTERCONNECT LAYERS OF $65nm$, $45nm$, $32nm$, $22nm$, AND OUR $16nm$ PROCESS TECHNOLOGIES.

Layer	Pitch (nm)				
	$65nm$ [15]	$45nm$ [16]	$32nm$ [10]	$22nm$ [13]	$16nm$
Metal 1	210	160	112.5	76	46
Metal 2	210	160	112.5	76	46
Metal 3	220	160	112.5	76	46
Metal 4	280	240	168.8	130	72
Metal 5	330	280	225.0	206	98
Metal 6	480	360	337.6	206	146
Metal 7	720	560	450.1	390	240
Metal 8	1080	810	566.5	390	240

TABLE II
DIMENSIONS OF METAL LAYERS USED IN OUR $16nm$ LIBRARY. ASPECT RATIO IS 1.9.

Layer	Width (nm)	Pitch (nm)	Thickness (nm)
Metal 1, 2, 3	22	46	41.8
Metal 4	32	72	60.8
Metal 5	44	98	83.6
Metal 6	66	146	125.4
Metal 7, 8	110	240	209
Metal 9, 10	400	800	760
Metal 11, 12	800	1600	1520

libraries. For the transistor model of our $16nm$ library, we use the PTM high-performance compact model for $16nm$ high-k metal gate and strained-Si CMOS ($16nm$ PTM HP model V2.1) [17].

C. Comparison of $45nm$, $22nm$, and $16nm$ Libraries

As mentioned earlier in this section, we design and analyze 2D ICs using commercial design and analysis tools with the $45nm$, the $22nm$, and our $16nm$ libraries, and compare wirelength, area, critical path delay, and power. We use Nangate $45nm$ standard cell library [14] and the $22nm$ standard cell library presented in [13] for $45nm$ and $22nm$ libraries, respectively.

1) Gate Delay: Gate delay and driving strength of a gate are determined primarily by transistor characteristics and the size of PMOS and NMOS in the gate. Therefore, the first experiment we perform is to compare the transistor characteristics. Especially, we focus on the strength of the transistors driving a load capacitor, so we perform SPICE simulations on the minimum-size inverter of each library.

Figure 2 shows the rise and the fall time of the minimum-size inverters. As the figure shows, the rise time of the $45nm$ transistor model is the largest, while the fall time of the $45nm$ transistor model is the smallest. This is due to unbalanced width ratio between the pull-up PMOS and the pull-down NMOS used in the minimum-size inverter. On the other hand, the rise and the fall time of our $16nm$ minimum-size inverter are smaller than those of the $22nm$ minimum-size inverter, when the load capacitance is small. However, when the load capacitance is large, the rise and the fall time of the $16nm$ minimum-size inverter become slightly larger than those of the $22nm$ minimum-size inverter. Although the transistor strength of our $16nm$ library is slightly weaker than that of the $22nm$ library, the actual gate delay of the $16nm$ library could be smaller than that of the $22nm$ library. This is mainly because the standard cells of the $16nm$ library are smaller than those of the $22nm$ library, so the input capacitance of the $16nm$ library is smaller.

We also compare the delay of minimum-size inverters driving other inverters. In this experiment, a minimum-size inverter in a process library drives another minimum-size inverter, which drives

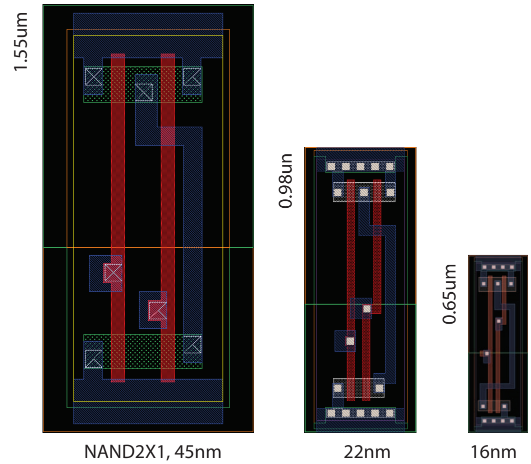

Fig. 1. $1\times$ two-input NAND gates in the $45nm$ [14], the $22nm$ [13], and our $16nm$ libraries (drawn to scale).

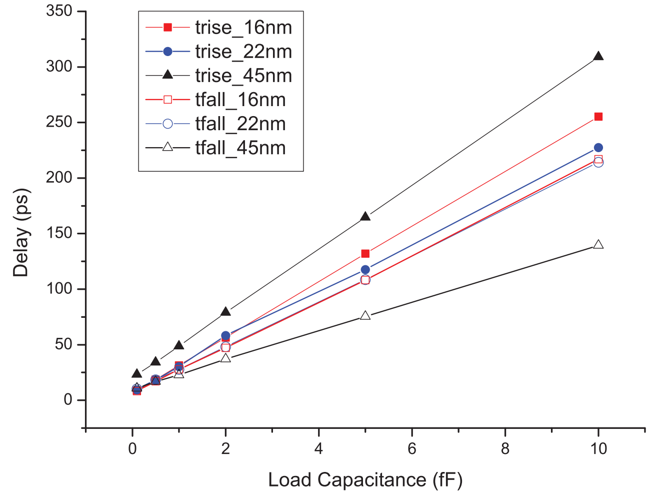

Fig. 2. Drive strengths of minimum-size inverters with a fixed load capacitance. RC parasitics are included.

an $N\times$ inverter in the same library. We obtain the delay of the second minimum-size inverter by SPICE simulation. Figure 3 shows the comparison results. We observe that the $16nm$ inverters have shorter delay than $22nm$ or $45nm$ inverters. Quantitatively, we observe approximately 30% improvement when the process moves from $45nm$ to $22nm$, and about 20% improvement when the process moves from $22nm$ to $16nm$. Notice that this SPICE simulation does not consider interconnect resistance and capacitance.

Since the input capacitance is an important factor determining delay and power, we show the input capacitances of $45nm$, $22nm$, and $16nm$ standard cells in Table III. As shown in the table, input capacitances of the $22nm$ standard cells are approximately half of the $45nm$ standard cells, and the $16nm$ standard cells have 5% to 20% smaller input capacitance than the $22nm$ standard cells.

2) Full-chip 2D Design: In this experiment, we design 2D circuits using the $45nm$, the $22nm$, and our $16nm$ libraries, and compare wirelength, area, critical path delay, and power. The experimental flow is as follows. We prepare two benchmark circuits shown in Table IV, synthesize, design, and optimize them using the $45nm$, the $22nm$, and the $16nm$ libraries. For all the libraries, we use the same area utilization for fair comparison, and we find the fastest

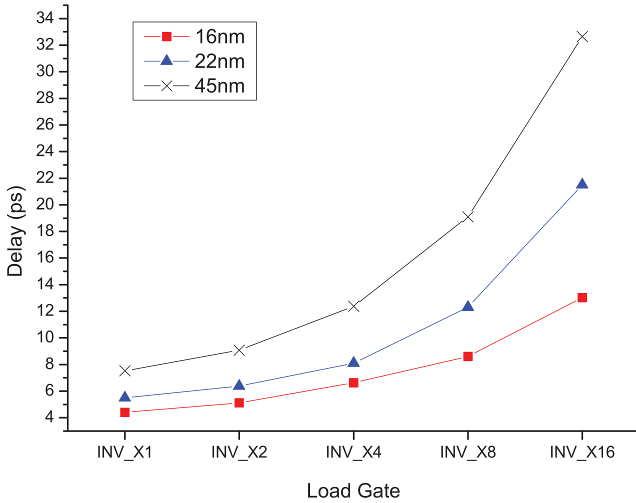

Fig. 3. Delay of a minimum-size inverter driving an N× inverter (N = 1, 2, 4, 8, 16), where both inverters are in the same process. RC parasitics are included.

TABLE III
INPUT CAPACITANCE OF SELECTED STANDARD CELLS IN THE $45nm$, THE $22nm$, AND THE $16nm$ LIBRARIES.

Cell	Cap (fF)		
	$45nm$	$22nm$	$16nm$
AND2 1×	0.54 (1.00)	0.25 (0.46)	0.22 (0.41)
AOI211 1×	0.64 (1.00)	0.30 (0.47)	0.25 (0.39)
AOI21 1×	0.55 (1.00)	0.23 (0.42)	0.20 (0.36)
BUF 4×	0.47 (1.00)	0.28 (0.60)	0.29 (0.62)
DFF 1×	0.90 (1.00)	0.41 (0.46)	0.26 (0.29)
FA 1×	2.46 (1.00)	1.31 (0.53)	1.36 (0.55)
INV 4×	1.45 (1.00)	0.69 (0.48)	0.56 (0.39)
MUX2 1×	0.95 (1.00)	0.42 (0.44)	0.34 (0.36)
NAND2 1×	0.50 (1.00)	0.24 (0.48)	0.22 (0.44)
OAI21 1×	0.53 (1.00)	0.25 (0.47)	0.20 (0.38)
OR2 1×	0.60 (1.00)	0.26 (0.43)	0.20 (0.33)
XOR2 1×	1.08 (1.00)	0.55 (0.51)	0.45 (0.42)
Average	(1.00)	(0.48)	(0.40)

operation frequency for each library. Table IV shows the number of gates, the number of nets, and the total cell area of the benchmark circuits.

Table V shows the comparison results for the 2D designs. In general, the chip area of the $45nm$ circuits is about three times larger than that of the $22nm$ circuits, and the chip area of the $22nm$ circuits is approximately two times larger than that of the $16nm$ circuits. In addition, the total wirelength of the $16nm$ circuits is approximately $1.48\times$ shorter than that of the $22nm$ circuits, and $3.08\times$ shorter than that of the $45nm$ circuits. Regarding the critical path delay, the $16nm$ circuits are $1.49\times$ faster than the $45nm$ circuits on average and $1.07\times$ faster than the $22nm$ circuits on average. Although the transistor strengths of the $22nm$ and the $16nm$ libraries are similar, the $16nm$ circuits have smaller critical path delay because the gate input capacitance is smaller. However, the wire resistance and capacitance of the $16nm$ library is slightly larger than that of the $22nm$ library, so we observe only 7% improvement in the critical path delay. Power consumption of the $16nm$ circuits is approximately $4.5\times$ smaller than that of the $45nm$ circuits and $1.1\times$ smaller than that of the $22nm$ circuits. Overall, the delay and power enhancement coming from $22nm$-to-$16nm$ transition is not as significant as the enhancement coming from $45nm$-to-$22nm$ transition because $45nm$

TABLE IV
STATISTICS OF OUR BENCHMARK CIRCUITS.

Circuit	# Gates	# Nets	Total cell area		
			$45nm$	$22nm$	$16nm$
BM1	$352K$	$372K$	0.632	0.218	0.098
BM2	$518K$	$680K$	1.288	0.437	0.198

TABLE V
COMPARISON OF 2D LAYOUTS.

	BM1 (350K gates)			BM2 (700K gates)		
	$45nm$	$22nm$	$16nm$	$45nm$	$22nm$	$16nm$
Area (mm^2)	1.00	0.36	0.17	2.56	0.81	0.42
Wirelength (m)	10.65	4.22	2.75	15.17	8.90	6.19
Delay (ns)	3.19	2.61	2.38	6.51	4.10	3.93
Power (W)	0.352	0.0684	0.068	0.521	0.154	0.133

and $22nm$ technologies are two generations apart, while $22nm$ and $16nm$ technologies are only one generation apart, and they use similar state-of-the-art structures, e.g., metal gate, low-k insulator, and high-k gate oxide.

III. FULL-CHIP 3D IC DESIGN AND ANALYSIS

A. Full-chip 3D IC Design and Analysis Methodologies

To design and analyze 3D ICs, we implement the 3D IC design and analysis methodologies [13]. We use the 3D placement engine obtained from [4], which implements a partitioning-based 3D global placement algorithm. We also vary the number of TSVs inserted in each design by exploiting different partitioning sequences. Detailed placement and routing on each die are performed by Cadence Encounter. Notice that for a fair comparison between 3D and 2D designs in different technologies, we set the area utilization to around 60% for all the designs. After generating 3D IC layouts, we perform 3D-aware timing analysis and optimization by 3D timing optimization tool [18].

B. Experimental Settings

The settings of our experiments are also extensions of the experiments in [13]. Since the number of gates in logic circuits is booming over the years, we use larger benchmark circuits, 256×8 Fast Fourier Transform and 256×16 Fast Fourier Transform. Statistics about these two benchmarks are shown in Table IV.

We use four dimensions of TSVs as shown in Table VI, which are the same as [13]. We use $5\mu m$ and $0.5\mu m$ TSVs in $45nm$ technology, and $1\mu m$ and $0.1\mu m$ TSVs in $22nm$ technology. $0.1\mu m$ TSVs with such high aspect ratio may be difficult to fabricate at present, but it provides an extreme case to explore the maximum benefits that 3D ICs may gain. In addition, it has been shown that the maximum TSV size combined with $22nm$ technology should be around $0.5\mu m$ [13]. Therefore, we use $0.5\mu m$ and $0.1\mu m$ TSVs in our $16nm$ library. The standard cell height of our $16nm$ technology is $0.6\mu m$. This means that a $0.5\mu m$ TSV occupies approximately two $16nm$ standard cell rows, while a $0.1\mu m$ TSV occupies one third of a $16nm$ standard cell row. Figure 4 shows GDSII images of TSVs and standard cells in $45nm$, $22nm$, and $16nm$ technologies.

IV. EXPERIMENTAL RESULTS

A. Comparison of $16nm$ 2D and 3D ICs

We first show the advantages of 3D ICs over 2D ICs in $16nm$ process by comparing four metrics (wirelength, footprint area, critical path delay, and power consumption) in Figure 5 and Figure 6. All the simulation results are normalized to the results of 2D circuits. In the

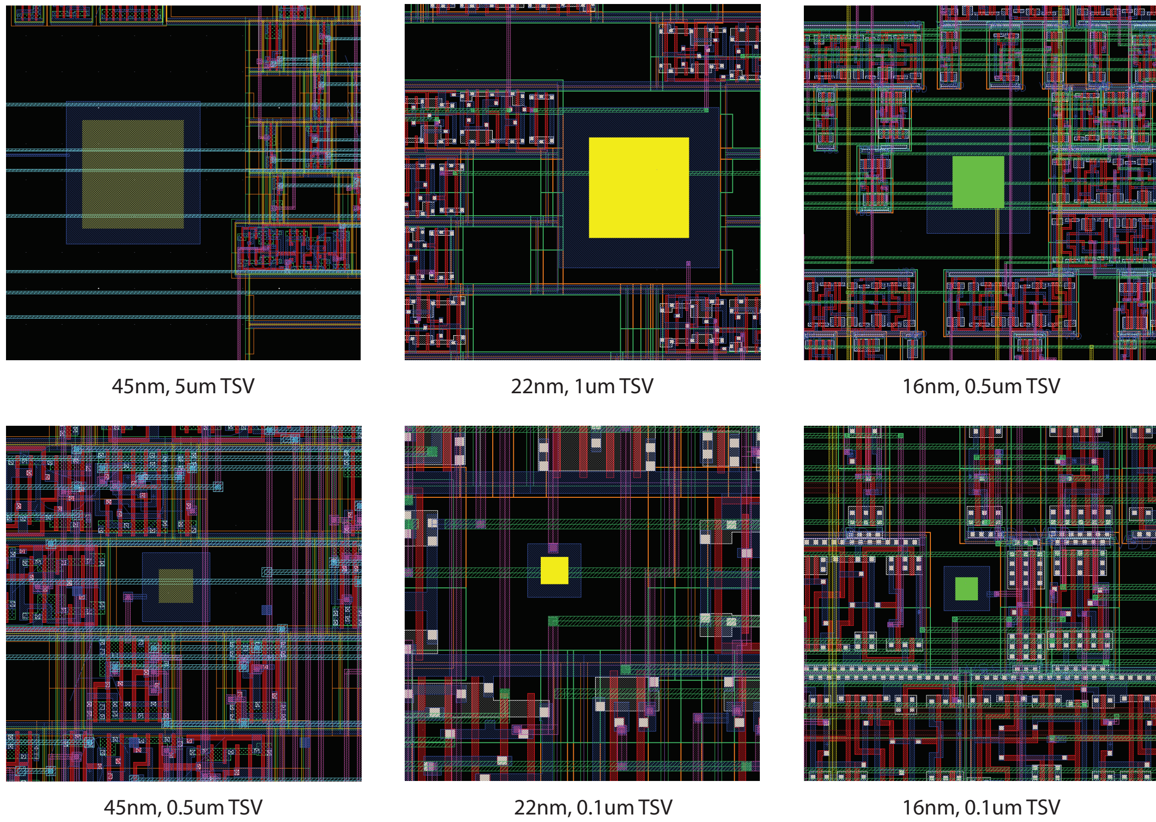

Fig. 4. Zoom-in GDSII layouts of the 6 types of designs studied in this paper. Each TSV is surrounded by its keep-out-zone.

TABLE VI
TSV-RELATED DIMENSIONS, DESIGN RULES, AND TSV CAPACITANCE.

Dimensions	TSV-5	TSV-0.5	TSV-1	TSV-0.1
Width (μm)	5	0.5	1	0.1
Height (μm)	25	8	5	5
Aspect ratio	5	16	5	50
Liner thickness (nm)	100	20	30	10
Barrier thickness (nm)	50	20	30	5
Landing pad width (μm)	6	1	1.6	0.18
TSV-to-TSV spacing (μm)	2	0.6	0.8	0.1
TSV-to-device spacing (μm)	1	0.36	0.4	0.1
TSV capacitance (fF)	20	3.2	2.67	0.8

experiments, we use three kinds of partitioning sequences during 3D placement to obtain different designs with various number of TSVs. If we apply z-direction cut first, we obtain fewer connections between two dies. On the other hand, if we apply z-direction cut later, we obtain much more vertical connections between two dies. Therefore, we generate min-cut designs, med-cut designs, and max-cut designs by applying z-direction cut at the beginning of the placement, in the middle of the placement, and at the last of the placement, respectively. The number of TSVs in each design is shown in Table VII.

1) Area and Wirelength: Figure 5 and Figure 6 show footprint area and wirelength results normalized to those of the $16nm$ 2D designs for BM1 and BM2. As shown in these figures, footprint area of the 3D designs is always smaller than that of the 2D designs by 20% to 50%. The total wirelength of the 3D designs is also smaller than that of the 2D designs by 9% to 18%. In addition, the min-cut designs have the smallest number of TSVs as shown in Table VII, so the area of the min-cut design is smaller than that of the med-cut and the max-cut designs.

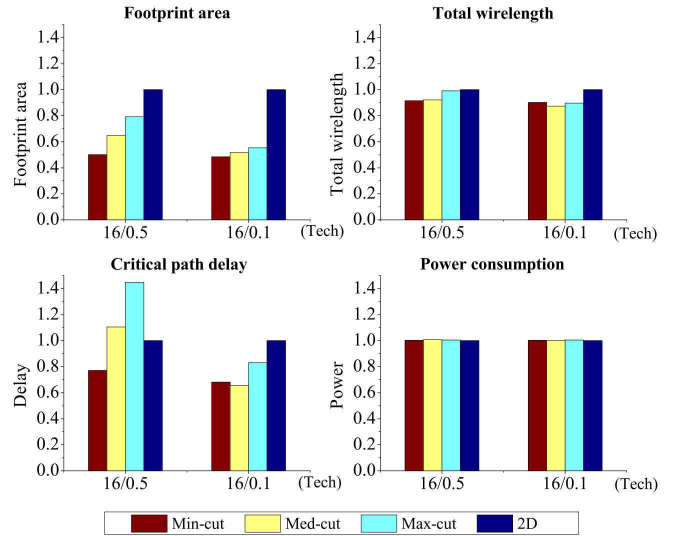

Fig. 5. Comparison of $16nm$ 2D and 3D designs with $0.5\mu m$ and $0.1\mu m$ TSVs for BM1.

It is obvious that wirelength depends on the area of 3D ICs, but there exist other factors affecting the total wirelength of 3D IC, e.g., placement, routing, TSV area, and so on as predicted in [2]. A noticeable point in the wirelength results is that the wirelength of the min-cut designs is smaller than that of the med-cut and the

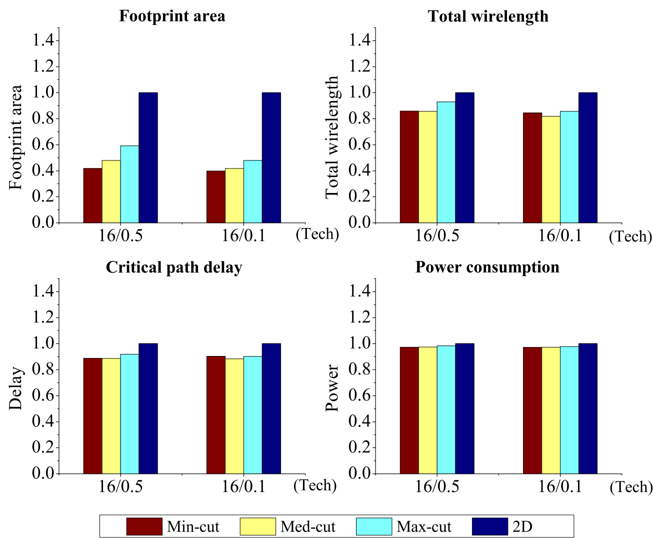

Fig. 6. Comparison of $16nm$ 2D and 3D designs with $0.5\mu m$ and $0.1\mu m$ TSVs for BM2.

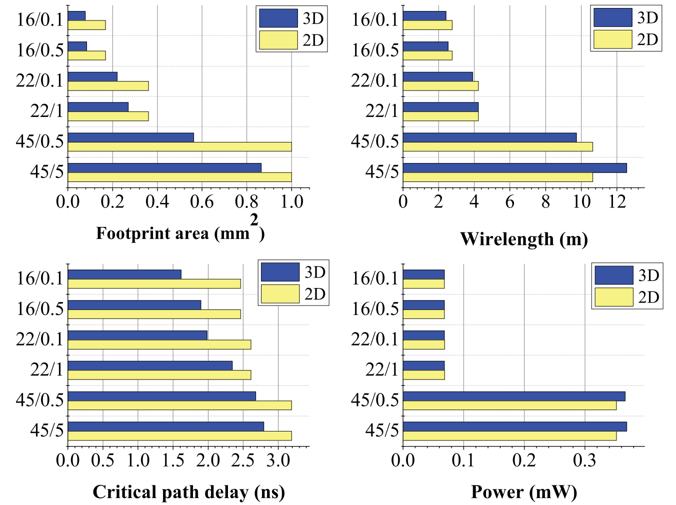

Fig. 7. Comparison of the optimized 2D and 3D designs (BM1) in $16nm$, $22nm$ and $45nm$ technology. The y-axis shows the technology combination (process node/TSV width).

TABLE VII
TSV COUNTS IN $16nm$ 3D DESIGNS.

Design	TSV width	Min-Cut	Med-Cut	Max-Cut
BM1	$0.5\mu m$	$1,337$	$8,776$	$13,967$
	$0.1\mu m$	624	$8,734$	$15,191$
BM2	$0.5\mu m$	$2,204$	$12,470$	$23,585$
	$0.1\mu m$	$2,204$	$12,470$	$23,586$

max-cut designs for $0.5\mu m$ TSV, while the wirelength of the med-cut designs is the shortest for $0.1\mu m$ TSV in some designs. A possible reason for this observation is that $0.5\mu m$ TSVs have a big impact on the area, so using fewer TSVs can reduce the wirelength. However, $0.1\mu m$ TSVs have a small impact on the area, so using more TSVs helps reduce wirelength.

2) Critical Path Delay: Figure 5 shows critical path delay of BM1 designed in 2D and 3D with $0.5\mu m$ and $0.1\mu m$ TSVs. In both $0.5\mu m$-TSV and $0.1\mu m$-TSV cases, the min-cut designs have the smallest critical path delay compared to the med-cut and the max-cut 3D designs and the 2D design. When $0.5\mu m$ TSVs are used, the min-cut 3D design shows about 20% improvement. Moreover, when $0.1\mu m$ TSVs are used, we observe about 40% improvement. However, the med-cut and the max-cut 3D designs using $0.5\mu m$ TSVs have larger cricital path delay than the 2D design. This result is due to several factors such as that the TSV capacitance is large, so using fewer TSVs is better than using more TSVs, and that the 3D placer we used is not a timing-driven placer but a wirelength-driven placer. On the other hand, all the 3D BM2 designs have smaller critical path delay than the 2D BM2 design as shown in Figure 6. In the figure, we observe about 10% improvement.

3) Power Consumption: As shown in Figure 5 and Figure 6, the power is almost the same in 2D and 3D designs regardless of the TSV size and the TSV count. As explained in [13], the reason for this observation is because shorter wirelength reduces the dynamic power consumption, but TSV capacitance increases it. In our experiments, the power of 3D BM2 designs has 5% reduction, while 3D BM1 designs do not show any power reduction. This is because BM1 and BM2 3D designs have similar wirelength reduction ratio (approximately 10%), but BM2 3D designs have smaller TSV count-to-wirelength reduction ratio than BM1 3D designs.

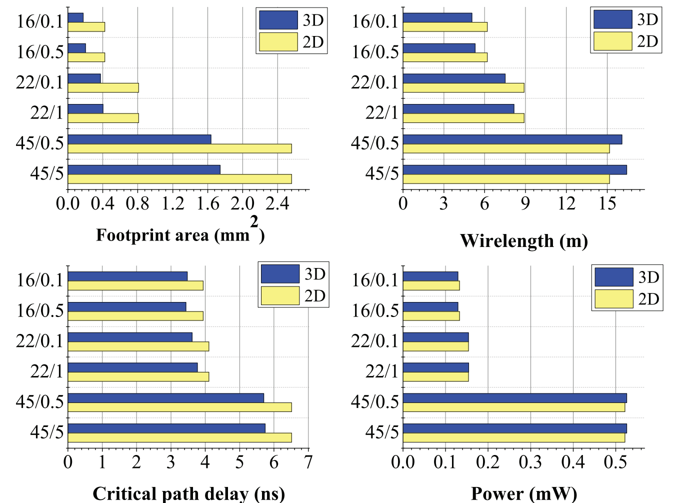

Fig. 8. Comparison of the optimized 2D and 3D designs (BM2) in $16nm$, $22nm$ and $45nm$ technologies. The y-axis shows the technology combination (process node/TSV width).

B. Comparison of 3D Designs in $16nm$, $22nm$ and $45nm$ Technology

Observing that there still exist benefits we can obtain from 3D ICs at $16nm$ technology node, we compare the performance of 2D and 3D ICs at different process nodes. To obtain the best-case results, we choose the best one from the min-cut, the med-cut, and the max-cut results in each technology and compare the four metrics we used in the previous sectinos. The results are shown in Figure 7 and Figure 8.

We observe that except the power consumption of $45nm + 5\mu m$ TSV cases, 3D ICs are better than 2D ICs. More importantly, $22nm$ 3D designs have similar footprint area and wirelength compared with the $16nm$ 2D designs, and the critical path delay of $22nm$ 3D designs are even smaller than that of $16nm$ 2D designs. This phenomenon is not observed between $45nm$ and $22nm$ designs because $45nm$ and $22nm$ are not consecutive process generations and have big disparity in gate performance.

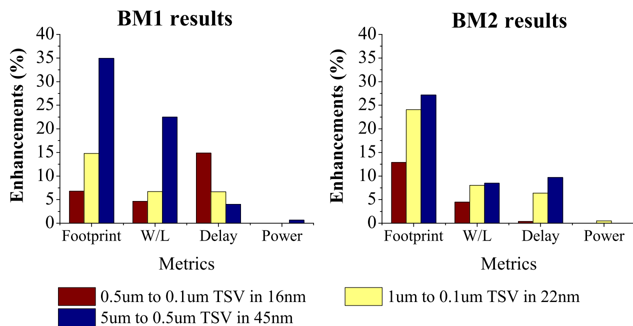

Fig. 9. Performance enhancement obtained by shrinking TSV size.

C. *The Impact of TSV Size*

Since TSVs have significant effects on die area, wirelength, critical path delay, and power, we vary the TSV size and study the impact of TSV dimension on the four metrics in this experiment. The results are shown in Figure 9. As shown in the figure, the footprint area is the most sensitive to the TSV size, and that the smaller footprint area achieved by using smaller TSVs is the direct benefit of shrinking the TSV size. Besides, wirelength is also directly related to the footprint area, so we observe wirelength reduction when we shrink the TSV size. However, delay improvement does not vary directly as the wirelength, because the critical path delay is sensitive to the TSV capacitance, gate delay, and wire delay. In addition, power remains almost unchanged in all the cases because the wire capacitance reduction is almost similar to the amount of TSV capacitance. More details are explained below.

Regarding the footprint area, the $45nm$ designs have larger enhancement than the $22nm$ and the $16nm$ designs when we shrink the TSV size. As to BM1, $5\mu m$ TSVs are too big for the $45nm$ process. Decreasing the TSV size by $10\times$ leads to a great benefit. On the other hand, $1\mu m$ TSVs and $0.5\mu m$ TSVs are already close to the proper TSV sizes for the $22nm$ and the $16nm$ processes. If the area percentage occupied by TSVs is already small, reducing the TSV size further does not lead to large chip area reduction. In fact, if we scale down the TSV size linearly with the device size, the TSV sizes for the $22nm$ and the $16nm$ processes should be $2.4\mu m$ and $1.8\mu m$, respectively, which are much bigger than the TSV sizes that we use. Table VIII shows the number of TSVs and area occupied by TSVs in each design. As shown in the table, the area occupied by TSVs is sufficiently small, when $1\mu m$ TSVs and $0.5\mu m$ TSVs are used with the $22nm$ and the $16nm$ processes, respectively.

Critical path delay of 2D ICs is mainly determined by placement, routing, and timing optimization. On the other hand, critical path delay of 3D ICs is also affected by the TSV parasitic capacitance. Therefore, although the wirelength reduction in the $45nm$ BM1 design is approximately 22%, the delay improvement is only 4%. On the other hand, the wirelength reduction in the $16nm$ BM1 design is small, but the delay improvement is about 15%, which primarily comes from very small TSV capacitance. Thus, we observe that TSV capacitance is as important as the wirelength with respect to the critical path delay. An important point to notice is that in some cases, critical paths are 2D paths. Table VIII shows the amount of TSVs in the critical paths. As seen in the table, three designs have no TSVs in their critical paths. However, we still obtain non-negligible enhancement when we shrink the TSV size. In this case, the delay improvement comes from the smaller footprint area and reduced wirelength.

TABLE VIII
ADDITIONAL TSV-RELATED STATISTICS. TSV AREA IS THE RATIO BETWEEN THE TOTAL AREA OCCUPIED BY TSVS AND THE TOTAL CHIP AREA. "C.P." DENOTES CRITICAL PATH.

	BM1					
	$45nm$		$22nm$		$16nm$	
TSV diameter	$5\mu m$	$0.5\mu m$	$1\mu m$	$0.1\mu m$	$0.5\mu m$	$0.1\mu m$
# TSVs	$1,029$	$17,385$	$1,757$	$8,718$	$1,337$	$8,734$
TSV area (%)	5.83	6.06	2.92	0.15	1.92	0.10
# TSVs in c.p.	1	0	3	4	2	4
	BM2					
	$45nm$		$22nm$		$16nm$	
TSV diameter	$5\mu m$	$0.5\mu m$	$1\mu m$	$0.1\mu m$	$0.5\mu m$	$0.1\mu m$
# TSVs	$1,959$	$13,316$	$2,700$	$13,056$	$12,470$	$12,470$
TSV area (%)	5.50	0.98	2.67	0.10	7.45	0.28
# TSVs in c.p.	0	0	1	2	1	2

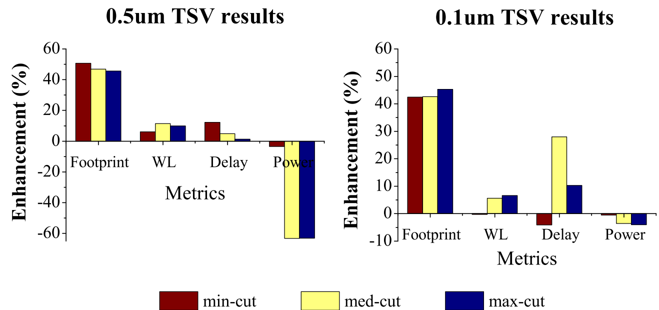

Fig. 10. Comparison between two-die and four-die 3D designs (BM1) built with our $16nm$ library.

D. *Comparison of $16nm$ Two-die and Four-die 3D ICs*

In this experiment, we build 3D ICs in two and four dies with our $16nm$ library and compare them in terms of footprint area, wirelength, critical path delay, and power. Figure 10 shows the comparison results (enhancements when we move from two-die designs to four-die designs). We observe that the four-die designs have 40% smaller footprint area, 6% to 11% shorter wirelength, and up to 27% better critical path delay than those of the two-die designs. However, the total silicon area increases by approximately 20% because four-die designs have more TSVs than two-die designs. Inserting more TSVs in the four-die designs also affects power consumption. As shown in Figure 10, the power increases when we stack more dies. This is due to the increased total capacitance induced by TSVs. Thus, we conclude that stacking more dies results in much smaller footprint area, shorter wirelength, and smaller critical path delay while suffering from power overheads because of the TSV capacitance.

V. CONCLUSIONS

In this paper, we investigated the performance of gate-level 3D designs in future technologies including the footprint area, total wirelength, critical path delay, and power consumption. We developed a $16nm$ technology standard cell library based on ITRS prediction and down-scaling trends of published process technologies. In addition, we explored the impacts of TSVs (dimensions and parasitic capacitance) on the quality of 3D ICs. The experimental results shows the following observations: 1) 3D ICs in $16nm$ process with nano-scale TSVs, if properly designed, achieve up to 40% reduction in footprint area and delay over 2D ICs; 2) 3D designs are comparable to 2D designs if their technology generaion gap is one or two; 3) by shrinking the TSV size to the nano scale, we obtain significant

benefits. We also recommended to use $0.5\mu m$ TSVs and $0.1\mu m$ TSV in $22nm$ and $16nm$ process nodes, respectively; 4) the four-die 3D ICs in $16nm$ technology have smaller footprint area, shorter wirelength, and better timing than the two-die 3D ICs, while suffering from power and silicon area overhead.

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, "TSV-aware Interconnect Length and Power Prediction for 3D Stacked ICs," in *Proc. IEEE Int. Interconnect Technology Conference*, June 2009, pp. 26–28.

[3] D. H. Kim, K. Athikulwongse, and S. K. Lim, "A Study of Through-Silicon-Via Impact on the 3D Stacked IC Layout," in *Proc. IEEE Int. Conf. on Computer-Aided Design*, Nov. 2009, pp. 674–680.

[4] M. Pathak, Y.-J. Lee, T. Moon, and S. K. Lim, "Through-Silicon-Via Management during 3D Physical Design: When to Add and How Many?" in *Proc. IEEE Int. Conf. on Computer-Aided Design*, Nov. 2010, pp. 387–394.

[5] M.-C. Tsai and T. Hwang, "A Study on the Trade-off among Wirelength, Number of TSV and Placement with Different Size of TSV," in *Proc. Int. Symp. on VLSI Design, Automation and Test*, Apr. 2011, pp. 1–4.

[6] D. H. Kim and S. K. Lim, "Through-Silicon-Via-aware Delay and Power Prediction Model for Buffered Interconnects in 3D ICs," in *Proc. ACM/IEEE International Workshop on System Level Interconnect Prediction*, June 2010, pp. 25–31.

[7] S. Gupta, M. Hilbert, S. Hong, and R. Patti, "Techniques for Producing 3D ICs with High-Density Interconnect," in *Proc. VLSI Multi-Level Interconnection Conf.*, 2004, pp. 56–59.

[8] M. Koyanagi, T. Fukushima, and T. Tanaka, "High-Density Through Silicon Vias for 3-D LSIs," in *Proceedings of the IEEE*, no. 1, Jan. 2009, pp. 49–59.

[9] ITRS, "International Technology Road Map for Semiconductors – Interconnect," http://www.itrs.net.

[10] P. Packan *et al.*, "High Performance $32nm$ Logic Technology Featuring 2nd Generation High-k + Metal Gate Transistors," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2009.

[11] Hou-Yu Chen *et al.*, "$16nm$ Functional $0.039\mu m^2$ 6T-SRAM Cell with Nano Injection Lithography, Nanowire Channel, and Full TiN Gate," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2009.

[12] D. H. Kim and S. K. Lim, "Impact of Through-Silicon-Via Scaling on the Wirelength Distribution of Current and Future 3D ICs," in *Proc. IEEE Int. Interconnect Technology Conference*, May 2011.

[13] D. H. Kim, S. Kim, and S. K. Lim, "Impact of Nano-scale Through-Silicon Vias on the Quality of Today and Future 3D IC Designs," in *Proc. ACM/IEEE International Workshop on System Level Interconnect Prediction*, June 2011.

[14] Nangate, "The Nangate $45nm$ Open Cell Library," http://www.nangate.com.

[15] P. Bai *et al.*, "A $65nm$ Logic Technology Featuring $35nm$ Gate Lengths, Enhanced Channel Strain, 8 Cu Interconnect Layers, Low-k ILD and $0.57\mu m^2$ SRAM Cell," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2004.

[16] K. Mistry *et al.*, "A $45nm$ Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, $193nm$ Dry Patterning, and 100% Pb-free Packaging," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2007.

[17] PTM, "Predictive Technology Model," http://ptm.asu.edu.

[18] Y.-J. Lee and S. K. Lim, "Timing Analysis and Optimization for 3D Stacked Multi-Core Microprocessors," in *IEEE Int. 3D System Integration Conf.*, Nov. 2010.