Logic-on-Logic 3D Integration and Placement

Thorlindur Thorolfsson□ , Guojie Luo† , Jason Cong† and Paul D. Franzon□
□ Department of Electrical & Computer Engineering, North Carolina State University, Raleigh, NC 27695
† Computer Science Department, University of California, Los Angeles, CA 90095
Email: gluo@cs.ucla.edu and cong@cs.ucla.edu

***Abstract*—In this paper we describe three 3D standard cell placement algorithms, which are: ``3D Placement using Sequential Off-the-Shelf 2D Placement Tools", ``True-3D Analytical Placement with mPL" and ``3D Placement using Simultaneous 2D Placements with mPL". We use these algorithms to place three case studies in a real face-to-face 3D integration process. The three case studies are a 2 point FFT butterfly processing element (PE), a Advanced Encryption Standard encryption block (AES) and a multiple-input and multiple-output wireless decoder (MIMO). The placements are then fully routed and compared to 2D placements in terms of performance and power consumption. Using this methodology we show that using 3D face-to-face integration with microbumps in conjunction with the three placement algorithms we can improve the maximum clock speed of AES module by 15.3% and the PE by 22.6%, while reducing the power of the AES module and the PE by 2.6% and 12.9% respectively.**

I. Introduction

This paper addresses the question as to what potential advantage of employing true 3D place and route tools in a synthesized logic-on-logic situation. In principal employing such tools should decrease the wiring length significantly. The reduction of on-chip wiring stemming from logic-on-logic 3D integration reduces power consumption and increases performance. The improvements in these metrics is the focus of this work.

In this work we analyze the power consumption and performance benefits of logic-on-logic 3D integration, through three case studies. These three case studies are a butterfly processing element (PE), an Advanced Encryption Standard (AES) module and a multiple-input and multiple-output wireless decoder (MIMO) wireless decoder. We believe these case studies represent a variety of design classes. The FFT butterfly processing element is a low power design with a very long critical path through a multiplier and two adders. The AES decryption module was obtained from OpenCores repository, and has much shorter critical path than the PE. This MIMO detector design is an implementation of a K-best sphere decoder that has a large number of flip-flops that are used for shift registers[4] The case studies are carried out in Tezzaron's 130 nm 3DIC process, which is described in Section III. For the case study the unit is placed using three different 3D placement algorithms and one 2D placement algorithm for comparison purposes. These placement algorithms are ``3D Placement using Sequential Off-the-Shelf 2D Placement Tools", ``True-3D Analytical Placement with mPL" and ``3D Placement using Simultaneous 2D Placements with mPL", which are described in Section IV-B, IV-C, and IV-D respectively.

II. Related Work

There has been some research work conducted in 3D standard cell placement. The most relevant of this research is the work by Hentschke et al.[2] and Deng et al.[3] In their work Hentschke et al. present a quadratic placement algorithm for placing standard cells that provides a 32% reduction in total wire length using 5 tiers on the ISPD 2004 benchmarks. Similarly, in their work Deng et al. show a 21.4% total wire length reduction for the golem3 benchmark using their 3D standard cell placer. Both these works focus on reducing total wire length, which is a very important metric. However, this metric does not present the complete picture as it is hard to directly translate into improvements in performance and power consumption. This work improves upon that by routing the placed circuits and directly analyzing the result in terms of power and performance.

III. 3D Integration Technology

The potential advantage of employing true 3D place and route tools to implement digital systems is highly dependent on the parameters of the 3D technology used. There are four main parameters that characterize a given 3D integration process and have the most impact on the results. These parameters are: the footprint of the via in micrometers, the minimum pitch at which two vias can placed next to each other (also in micrometers), which metal layers are blocked by placing the via and whether a given via has to be placed on a grid or can be placed freely. Table I shows the four parameters of for 3D vias that are a part of Tezzaron, MIT Lincoln Laboratory 3D integration process [4] and lower cost through-silicon via (TSV).

TABLE I
PARAMETERS OF 3D VIAS.

Via	Footprint	Pitch	Block	Grid
MIT Laser-Drilled TSV	2.5 by 2.5	3.9	All	No
Tezzaron Super-contact TSV	1.2 by 1.2	1.76	All	No
Tezzaron Copper Microbump	4.4 by 4.4	5.0	None	Yes
Low Cost TSV	5.0 by 5.0	5.0	All	No

We use the technology parameters of Tezzaron's[5],[6],[7] 130 nm 3D technology process for all the placements. The advantage of this process is that through-tier connections do not block any routing because the process uses microbumps in face-to-face (as shown on the left of Figure 1 configuration instead of TSVs to communicate between unavoidable.

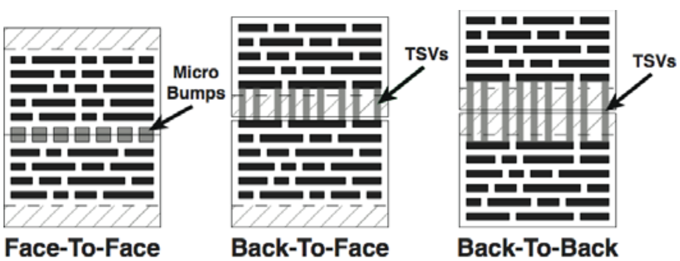

Figure 1. The three different stacking orientations, with the interconnect and substrate shown.

IV. 3D PLACEMENT

This section describes the three 3D placement approaches, which are: ``3D Placement using Sequential Off-the-Shelf 2D Placement Tools", ``True-3D Analytical Placement with mPL" and ``3D Placement using Simultaneous 2D Placements with mPL", along with 2D placement that is used for comparison.

A. 2D Placement using Off-the-Shelf Tools

2D placement using off-the-shelf tools is as the name implies a traditional 2D placement that is done using commercial tools. The commercial tool that are used for both placement and routing in this case is Cadence Encounter. This is very convenient as the same tool is also used for routing the other placements and as such insures a fair comparison between 2D and 3D placements.

B. 3D Placement using Sequential Off-the-Shelf 2D Placement Tools

The first placement algorithm we explore is the one that is the most similar to 2D placement in the sense that it basically realizes a 3D placement from two separate 2D placements and it uses off-off-shelf commercial tools to obtain its placement. It works in the following manner. First, the netlist of the design is represented by a hypergraph where the nodes correspond to the standard cells and the vertices correspond to the nets that connect the standard cells. The hypergraph is then partitioned into into two balanced halves that have as few edges between the halves as possible. This partitioning is done using hMetis[8] and in the partitioning the area balance is favored over minimizing the number of edges. This is done to ensure that area is not wasted due to an imbalance between the tiers. Placement is then completed using Cadence Encounter as follows. First, the first tier is 2D placed without any constraints from either the input or the output pins. This placement is then used for 3D via assignment as described in Section V. The 3D via assignment is then used to constrain a re-placement of the first tier, followed by a placement of the second tier.

C. True-3D Analytical Placement with mPL

This placement is implemented by an analytical 3D placer, mPL-3D[9] This analytical placer formulates and solves the 3D placement as a nonlinear programming problem (NLP). The problem variables include the horizontal placement (x,y) of each cell, and the vertical placement (tier assignment) z. The intermediate tier assignment is relaxed to allow fractional placement between two neighboring tiers, and it is eventually legalized when the problem is solved. The NLP has an objective as the weighted sum of half-perimeter wirelength (HPWL) and the number of 3D vias, with the constraints that the bin-wise area of standard cells is less than the bin capacity in a binned 3D placement region. The constraints are converted to a penalty term as the sum of squares of bin overflows, and a legalized solution is obtained when a sequence of penalized objectives are solved with an increasing penalty factor. For the placements of the case studies, we set the weight of 3D vias in the objective function with a small number (0.1), which result in a high number of 3D vias. Thus, it tries its best to minimize the HPWL by allowing a certain amount of nets across tiers.

D. 3D Placement using Simultaneous 2D Placements with mPL

This pseudo-3D placement is also implemented by mPL-3D. It performs in a way that is a mix between between the placements in Section IV-B and IV-C. It uses fewer vias than the placement in Section IV-C In this placement algorithm the tier assignment is done by using hMetis just like in Section IV-B. However, instead of performing two separate 2D placements, it calls mPL-3D with the tier assignment z fixed, and only has the horizontal placement (x,y) as variables. This improves upon the placement in Section IV-B as each tier directly influences the placement on the other tier.

V. 3D ROUTING

In order to analyze the power and performance benefits of logic-on-logic 3D integration, it is necessary to complete 3D routing of the placed cells. Luckily, there is less difference between 2D and 3D routing than between 2D and 3D placement because the multiple metal layers used for 3D routing create a 3D structure that is similar to the one that is used for 2D wiring. Although, there has been some research effort spent on native 3D routing [10],[11]. The approach we use in this work involves decomposing the 3D routing problem into separate 2D routing problems that can be solved with conventional 2D routing tools. We do this by employing a 3D via assignment algorithm that is explained in Section V-A.

A. 3D Via Assignment

This section describes the 3D via assignment algorithm. The algorithm was originally proposed by Thorolfsson et al.[12] and was used for all three of the 3D placement algorithms. The algorithm is based on Lee's algorithm[13] and works in the following manner. First, a grid is generated that corresponds to the 3D via grid. Each wire that travels between tiers is then assigned to the grid square that is the closest to its placement location. This results in some grid squares having multiple wires assigned to them. After the initial assignment, a shifting operation is performed on every grid square that has more than one inter-tier wire. The shifting operation starts with the grid squares that have the highest number of inter-tier wires and proceeds downward. The shifting operation works in the following manner. The shortest path from the grid square to a free grid square is found using Lee's algorithm. Along that path the content of every grid square along that path is shifted one grid square towards the free square. This reduces the number of inter-tier wires in the target square by one. The shifting operation is performed until no square has more than one inter-tier wire. The algorithm is shown below:

```
Input: Location of cells that connect to 3D vias
Output: The 3D via assignment
AssignEveryInterTierSignalToNearestGridSquare();
foreach Grid Square i j do
    if 3D vias assigned to i j > 1 then
        while 3D vias assigned to i j > 1 do
            k = ShortestPathToFreeGridSquare();
            foreach 3D Via on path k do
                Shift3DViaAlongPath();
            end
        end
    end
end
```

VI. Results

The results of placing the three case studies are generated in the following manner. First, each of the case studies is synthesized to a gate level netlist using Synopsys Design Compiler. This gate level netlist is then is placed in Cadence Encounter using either Encounter's placer or mPL-3D's placer depending on the placement algorithm used. The number of 3D vias is different for each placement algorithm. The True-3D Analytical Placement algorithm uses more 3D vias than the other two placement algorithms. Table II shows the utilization of the placements for each case study.

TABLE II
THE 3D VIA UTILIZATION NUMBERS THE 3D PLACEMENT ALGORITHMS FOR THE DIFFERENT CASE STUDIES.

	3D Vias Used	3D Vias Available)	Util.
PE 3D Seq.	362	6241	5.8%
PE 3D Sim.	338	6400	5.3%
PE 3D True	1335	6400	20.9
AES 3D Seq.	240	4900	4.9%
MIMO 3D Seq.	2698	8218	32%

For each of the placements, the cells that use the clock signal are kept on the same tier. Keeping these cells in the same tier makes the design more resistant to process variation as all the clock buffers are manufactured on the same wafer. After placement, clock tree synthesis is performed using Cadence Encounter Cts. in the tier that has the clock, followed by routing of both tiers. After routing, the interconnect parasitics are extracted into a SPEF file and the total wire length of the route is calculated. The SPEF file is then read into Synopsys PrimeTime. In PrimeTime the maximum clock period along with the power consumption is determined from the parasitics of the SPEF file. Due to time limitations, the Sequential and the True 3D placements were not completed. The results for the placements are presented numerically in the Table III, followed by the percentage improvement over 2D in Table IV. Additionally, a visualization of the placed and routed AES case study for the 2D placement, and the top of bottom of the sequential 3D placement is shown in Figures 2, 3 and 4 respectively.

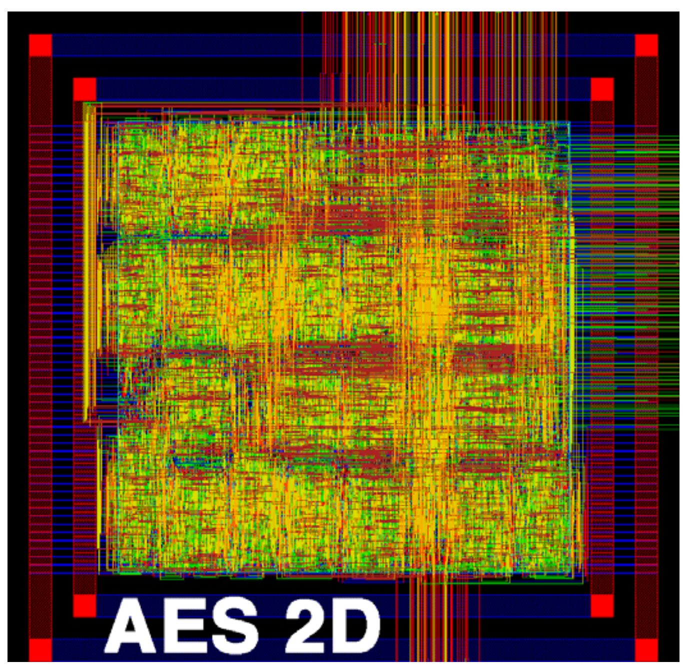

Figure 2. A picture of the 2D placed AES module.

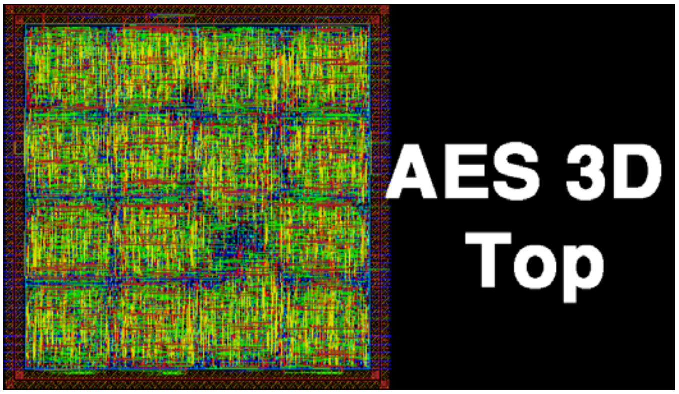

Figure 3. A picture of the top of the 3D sequentially placed AES module.

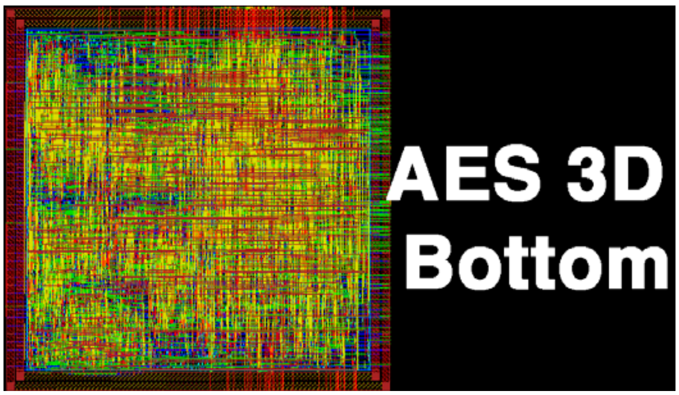

Figure 4. A picture of the bottom of the 3D sequentially placed AES module.

VII. CONCLUSION

We have shown that using logic-on-logic integration can improve the maximum clock frequency (up to 22.6%) and the power consumption (up to 12.9%) of different digital circuits. These improvements are realizable with little additional work using current 2D tools as we demonstrated with the 3D Placement using Sequential Off-the-Shelf 2D Placement Tool Algorithm, but are definitely more substantial when true 3D placement is used. Furthermore, the use of face-to-face integration does improve the result because the microbumps, unlike TSVs do not block any routing. Finally, these results are contingent on having enough 3D connections available for a given number of cells. If the cell size is reduced via a process shrink without an accompanied reduction in the the 3D via pitch the benefits of logic-on-logic may not be as substantial.

ACKNOWLEDGMENT

This work was supported by the MARCO Musyc Center, by DARPA under contract FA8650-04-C-7127 and contract FA8650-04-C-7120 and by Semiconductor Research Corporation. The authors thank Synopsys, Mentor Graphic and Cadence for the use of their University tools in this work. The authors would also like to thank Steve Lipa.

REFERENCES

[1] N. Moezzi-Madani, T. Thorolfsson, and W. Davis, ``A low-area flexible mimo detector for wifi/wimax standards," mar. 2010, pp. 1633 --1636.

[2] R. Hentschke, G. Flach, F. Pinto, and R. Reis, ``Quadratic placement for 3d circuits using z-cell shifting, 3d iterative refinement and simulated annealing," in SBCCI '06: Proceedings of the 19th annual symposium on Integrated circuits and systems design New York, NY, USA: ACM, 2006, pp. 220--225.

[3] Y. Deng and W. P. Maly, ``Interconnect characteristics of 2.5-d system integration scheme," in ISPD '01: Proceedings of the 2001 international symposium on Physical design New York, NY, USA: ACM, 2001, pp. 171--175.

[4] V. Suntharalingam, R. Berger, J. Burns, C. Chen, C. Keast, J. Knecht, R. Lambert, K. Newcomb, D. O'Mara, D. Rathman, D. Shaver, A. Soares, C. Stevenson, B. Tyrrell, K. Warner, B. Wheeler, D.-R. Yost, and D. Young, ``Megapixel cmos image sensor fabricated in three-dimensional integrated circuit technology," Solid-State Circuits Conference, 2005. Digest of Technical Papers. ISSCC. 2005 IEEE International, pp. 356--357 Vol. 1, Feb. 2005.

[5] R. Patti, ``Three-dimensional integrated circuits and the future of system-on-chip designs," Proceedings of the IEEE, vol. 94, no. 6, pp. 1214--1224, June 2006.

[6] ------, ``Interlocking conductor method for bonding wafers to produce stacked integrated circuits," U.S. Patent 6 838 774, January 4 2005.

[7] Tezzaron. Wafer stack with super-contacts. [Online]. Available: http://www.tezzaron.com/about/PhotoAlbum/Products/Wafer_Pair_Super-Contacts.html

[8] G. Karypis and V. Kumar, ``Multilevel k-way hypergraph partitioning," in Design Automation Conference, 1999. Proceedings. 36th}, 1999, pp. 343--348.

[9] J. Cong and G. Luo, ``A multilevel analytical placement for 3d ics," jan. 2009, pp. 361 --366.

[10] R. Enbody, G. Kwee, and H. Tan, ``Routing the 3-d chip," 1991, pp. 132 --137.

[11] C. C. Tong and C.-L. Wu, ``Routing in a three-dimensional chip, "Computers, IEEE Transactions on}, vol. 44, no. 1, pp. 106 --117, jan. 1995.

[12] T. Thorolfsson, N. Moezzi-Madani, and P. D. Franzon, ``Reconfigurable five layer 3d integrated memory-on-logic synthetic aperture radar processor,"To appear in Computers Digital Techniques, IET, vol. 4, no. 6, dec. 2010.

[13] C. Lee, ``{An algorithm for path connections and its applications}," IRE Transactions on Electronic Computers}, vol. 10, no. 2, pp. 346--365, 1961.

TABLE III

RESULTS FROM THE FOUR PLACEMENT ALGORITHMS, WITH ALL THE POWER MEASUREMENT DONE AT THE MAXIMUM CLOCK SPEED OF THE 2D PLACEMENT.

	Total Wire Length (mm)	Max Frequency (MHz)	Parasitic Power (mW)	Total Power (mW)
PE 2D	588.0	31.61	1.794	5.975
PE 3D Seq.	487.3	33.84	1.516	5.692
PE 3D Sim.	484.1	36.72	1.293	5.515
PE 3D True	464.8	38.74	0.984	5.206
AES 2D	460.9	250.63	5.1	38.0
AES 3D Seq.	423.9	289.02	4.1	37.0
MIMO 2D	307.8	221.23	6.3	43.2
MIMO 3D Seq.	973.0	259.07	4.1	41.0

TABLE IV

RESULTS FROM THE FOUR PLACEMENT ALGORITHMS, WITH ALL THE POWER MEASUREMENT DONE AT THE MAXIMUM CLOCK SPEED OF THE 2D PLACEMENT.

	Total Wire Length (% Change)	Max Frequency (% Change)	Parasitic Power (% Change)	Total Power) (% Change)
PE 3D Seq.	-17.1%	+7.1%	-15.5%	-4.7%
PE 3D Sim.	-17.7%	+16.2%	-27.9%	-7.7%
PE 3D True	-21.0%	+22.6%	-45.2%	-12.9%
AES 3D Seq.	-8.0%	+15.3%	-19.6%	-2.6%
MIMO 3D Seq.	+216.1%	+17.1%	-34.9%	-5.1%