Optimal Number and Placement of vertical links in 3D Network-on-Chip

Nesrine Toubaline, Djamel Bennouar, Ali Mahdoum

***Abstract*—**3D technology can lead to a significant reduction in power and average hop-count in Networks on Chip (NoCs). It offers short and fast vertical links which copes with the long wire problem in 2D NoCs. This work proposes heuristic-based method to optimize number and placement of vertical links to achieve specified performance goals. Experiments show that significant improvement can be achieved by using a specific number of vertical interconnect.

***Keywords*—** Interconnect optimization, Monolithic Inter-tier Vias, Network on chip, System on chip, Through Silicon Vias, Three Dimensional Integration Circuits.

I. Introduction

ACCORDING to the International Technology Roadmap for Semiconductors (ITRS), the number of transistors integrated on a chip continues to grow [1]. Three dimensional integrated circuits (3D-ICs) have emerged as a promising solution to continue the progress of Moore's law. 3D technology allows in reducing the problem of long interconnects lengths in 2D ICs by stacking multiple dies vertically. The major advantage of 3-D ICs is the considerable reduction in the length and number of global interconnects, resulting in an increase in the performance and decrease in the power consumption and area of wire limited circuits [2], [3]. It is shown that a 3D architecture can reduce wire length as much as the square root of the number of stacked layers [4]. They can be realized using Through Silicon Vias (TSVs), or Monolithic Inter-tier vias (MIVs). Recent researches have focused on TSV [5]. Compared with traditional wire bonding 3D integration, the TSV is significantly shorter and thus have a reduced delay and lower power consumption [6]. The TSV manufacturing yield and cost for 3D IC are studied in [7]. Equations for determining the TSV manufacturing costs per number of TSVs have been provided. As the number of TSVs grows, e.g. 10,000 TSVs per chip, the probability of a faulty chip die becomes higher [8]. It is obvious that the maximum performance can be achieved by full layer-layer connection, however, as the number of tiles grow, it might not be practical to assume that each tile will be connected with corresponding TSVs because of the limitation of manufacturing cost and chip area [9]. The number of TSVs in a 3D chip should be controlled, so that the manufacturing cost of a 3D chip is acceptable by the market [10]. An emerging alternative to TSV-based 3D is monolithic 3D that enables orders of magnitude higher integration density compared to that of TSV-based technology, due to the extremely small size of the monolithic inter-tier vias (MIVs) [11]. MIVs provide better electrical characteristics (i.e., less parasitics, electrical coupling, etc.) than TSVs, and also enable higher integration densities due to their small size [11]. Since the MIV size is negligible, reduce the number of MIVs, opening up the possibility for further optimization. In this article, we address a greedy heuristic to optimize number and placement of TSVs/MIVs (vertical links) in 3D network on chip.

The remainder of this paper is organized as follow. Section 2 presents related work. In section 3, we propose our heuristic approach. Section 4 reports our experimental results. Finally, in section 5, we make some concluding remarks.

II. RELATED WORK

In recent years, a number of research efforts have addressed the 3D integration technology. Three dimensional integrated circuits can be realized using vertical interconnects (Through Silicon Vias -TSVs, or Monolithic Inter-tier vias -MIVs). Length, delay and power consumption between vertical and horizontal interconnects are typically asymmetric in nature [12], vertical interconnects outperforming horizontal ones [13]. Compared to TSV-based 3D integration, monolithic integration can lead to up to 8% smaller area due to smaller size of inter-tier via than the TSVs, 12% lower longest path delay, and 7% lower power [14].

In [9], placements of full, half and quarter TSV number are compared for 4*4 3D mesh of 2 layers using hop count metric. They proceed by exhaustively enumerate all placement possibilities and output the combinations of lowest hop count. Nevertheless exhaustively enumerate is not possible for a large search space. In [13], a combined solution to the TSV placement and mapping of cores to routers in a three-dimensional Network-on-Chip (NoC) design is proposed. They include a constraint on the placement of the TSVs that no two TSVs can be within one hop distance of each other (minimum of two hop distance between the TSVs while placing them). In the two last works a fixed number of TSV is required. An integer linear programming (ILP) model is proposed in [15] to minimize the number of through-silicon-vias (TSVs) under both the

N. Toubaline is with the Department of Computer science, University Saad Dahleb of Blida, Blida, Algeria (Corresponding author e-mail: toubaline_n@hotmail.com).

D. Bennouar is with the Department of Computer science, University of Bouira, Bouira, Algeria (e-mail: dbennouar@ens-kouba.dz).

A. Mahdoum is with the Division of Microelectronics and Nanotechnologies, centre de Développement des Technologies Avancées, Alger, Algeria.

layer number constraint and the footprint area constraint; however the placement of TSVs is not discussed.

In this paper, we consider both optimizations the placement and number of vertical links. The novel contributions of the paper are: (i) obtaining the optimum placement of vertical links by considering the set of transfers between the IPs (shortest path). (ii) For the obtained placement of vertical links, found the optimal number of vertical links in order to minimize the hop count. The next section explains our heuristic.

III. THE PROPOSED HEURISTIC

The problem that we are trying to solve can then be stated as follows: Given the placement of the set of IPs (blocks) in 3D topology, the set of transfer between IPs, determine the number and location of all TSVs/MIVs that aims to finding a balance between performances (minimize delay, power and area).

It is obvious that the minimal delay can be achieved by full layer-layer connection; however, this will be to the detriment of area. To find the optimal solution, we can tray all possible combination of number and placement of TSVs/MIVs and take the best combination. It is noteworthy that exhaustive simulation is feasible for a small search space, while a heuristic-guided search is required to deal with the computational complexity from a larger search space [16]. In our paper, the hop count is selected as the metric in evaluating the performance of a NoC.

The inputs of the problem are the IPs configuration and a set of transfer between these IPs. The proposed idea is to use a greedy strategy: TSVs/MIVs are inserted one by one until we find a solution that satisfies the constraint (*less than a max cost*). A Hop Count is selected as a metric in evaluating the performance of a configuration. The outputs of this heuristic are: the number and location of all TSVs or MIVs (we note VL vertical link to express MIV or TSV). Below is the algorithm of our heuristic-based method:

A. Algorithm

1. Determine the optimal placement of VLs candidates (*the max number of VLs*).
2. Calculate the cost of all transfers against each VL candidate (*even assuming that VL can be used for any transfer*).
3. Find the optimal solution that minimizes the total cost (*from combinations of possible transfer paths*).

B. Explanation

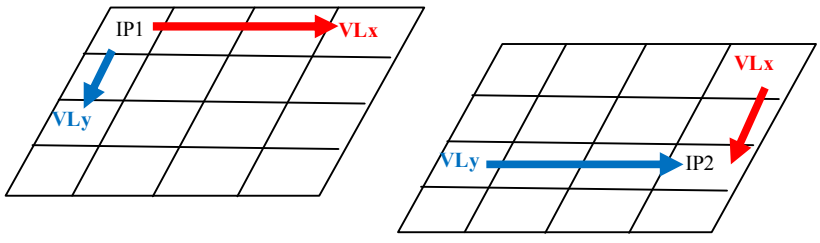
1. **Determine the optimal placement of VLs candidates**: the number of possible placement of VLs determines the maximum number of VLs candidates. We obtain the optimum placement of vertical links by considering the shortest path of each transfer between two IPs. It was assumed initially insert two positions (VLs) to each transfer between two layers (VLx and VLy). So we will have the maximum number of candidates VLs = 2 * number of transfers. The VLx and VLy positions are calculated from the IPs positions of a transfer.

Example: Let a transfer between IP1 (position-i, position-j) to the layer 1 and IP2 (position-i, position-j) of the layer 2 (fig. 1): the position of VLx = (position-i IP1, position -j IP2) and the position of VLy = (i-position IP2, IP1 position-j). The VLs candidates determine the possible ways to make a transfer between any two IPs.

2. **Calculate the cost of all transfers against each VL candidate**: Knowing that VLs candidates determined in the previous phase can be used for any transfer. Step 2 is to estimate the cost of each transfer with respect to each candidate VL.

TABLE I: TABLE OF PATHS (EACH PATH IS CHARACTERIZED BY A COST)

	VL1x	VL1y	….	VL*X*y
Transfer 1	Cost	…		
…	….			
Transfer *X*				



Layer 1 Layer 2

Fig. 1 Positions of VLs candidats

3. **Seek the optimum solution**: For each combination of paths from the previous step, calculate the total cost of transfers. The optimal solution is the one that gives the minimum cost. For N transfers, there will be $(2*N)^N$ possible paths combinations to calculate.

For a large number of possible combinations, the following heuristic is proposed; the idea is to use a greedy strategy: VLs candidates are inserted one by one until we find a solution that satisfies the constraint (less than a max cost).

Inserted VLs are sorted in ascending order of their transfer costs calculated from the previous step (the total cost of a VL is the sum of the costs of each transfer). Each time we insert a new VL, the candidate is assigned to a set of transfers (*the choice of transfers and the number of transfer is random*). The assignment is performed k times to change the assignment of the various candidate VL transfers (*k parameter is also random*). We take the solution (assignment) that minimizes the total cost.

IV. EXPERIMENTAL RESULTS

In this section, we present the experimental evaluation under different number of transfer (9 benchmarks) between 32 IPs in 2*8*2 3D architecture (2 lines, 8 columns and 2 layers). We applied the exact method (exhaustive enumeration) for the small size of research space to compare with our heuristic. Table II shows the experimental results.

The column #transfer shows the number of transfer between IPs (for N transfers the size of research space is $(2*N)^N$), the column #VL denotes the number of proposed VLs, for the columns cost, we calculate the minimum cost when using just one VL (minimum number of VLs); the exact method calculates the minimum cost (all possible combination of number of VLs used by all transfers). For the heuristic, cost constraint (< cost of using one VL) can be specified by the user when the cost of using one VL is not appreciated.

The exact method proves that a maximum number of LVs is equal to number of transfers. Remember that we have taken the shortest path (minimum hop count) to fixe LVs placement. Compared with a full layer-layer connection that requires 16 LVs (for a 2*8 NoC), our technique requires not even half VLs for all benchmarks (for 2 to 6 VLs). Assigned that 5 different LVs are proposed to 5 transfers, the 11 LVs are unused in full connection that will be in detriment of area. Note that our approach guarantees minimizing the number of LVs.

Minimize delay and area are two conflicting constraints. We calculate the cost (the 3rd column of table II) of minimum area (with 1 LV) and propose to introduce the cost constrain by the designer to improve delay (hop count).

V. CONCLUSION

In this paper, we have proposed optimal number and placement of vertical interconnects (TSVs or MIVs) in 3D networks on chip. The first phase involves determining the candidates' positions of vertical interconnect using the shortest path between IPs of each transfer. Then, the second phase to reduce the number of vertical interconnects. The optimum number of VLS was determined by greedy heuristic under hop count metric. The results are compared for exact method, showed that our heuristic can give the same improvement as the exact method.

REFERENCES

[1] S.I. Association, The international technology roadmap for semiconductors (ITRS), http://www.itrs.net/Links/2011ITRS/Home2011.htm (2011)

[2] J. W. Joyner *et al.*, Impact of three-dimensional architectures on interconnects in gigascale integration, IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 9, no. 6, pp. 922–927, Dec. 2000.

[3] V. F. Pavlidis and E. G. Friedman, Interconnect delay minimization through interlayer via placement in 3-D ICs, in Proc. ACM Great Lakes Symp.VLSI, 2005, pp. 20–25

[4] J. Joyner, P. Zarkesh-Ha, and J. Meindl, A stochastic global net-length distribution for a three-dimensionalsystem-on-a-chip (3d-soc), in Proc.of the 14th IEEE Int. ASIC/SOC Conf., September 2001, pp. 147–151.

[5] V. Pavlidis and E. Friedman, Three-Dimensional Integrated Circuit Design. Morgan Kaufmann, 2008.

[6] D. Velenis et al., Impact of 3D design choices on manufacturing cost, in: IEEE 3DIC 2009, 2009, pp. 1–5.

[7] J.H. Lau, TSV manufacturing yield and hidden costs for 3D IC integration, in: Electronic Components and Technology Conference (ECTC), 2010 Proceedings 60th, 2010, pp. 1031–1042

[8] C.-C. Chan et al., 3DICE: 3D IC cost evaluation based on fast tier number estimation, in: Proc. of of the 12th ISQED, 2011, pp. 1–6.

[9] T. C. Xu, P. Liljeberg, H. Tenhunen, A study of Through Silicon Via impact to 3D Network-on-Chip design, 2010 International Conference on Electronics and Information Engineering (ICEIE 2010) volume 1 –pages 333-337

[10] T. C. Xu, G. Schley, P. Liljeberg, M. Radetzki, J. Plosila, H. Tenhunen, Optimal placement of vertical connections in 3D Network-on-Chip. Journal of Systems Architecture 2013

[11] S. Panth, K. Samadi, Y. Du, and S. K. Lim, High-Density Integration of Functional Modules Using Monolithic 3D-IC Technology. Design Automation Conference (ASP-DAC), 18th Asia

TABLE II: EXPERIMENTAL RESULTS

Benchmark	#transfer	Cost (of 1 VL)	Exact method		Our heuristic		
			Cost	#VL	Cost constraint	Cost	#VL
1	2	17	9	2	15	9	2
2	3	25	11	2	24	21	2
					20	11	3
					15	11	3
3	4	35	11	3	30	25	2
					20	19	3
					12	11	3
4	5	42	10	5	35	34	2
					25	20	4
					11	10	5
5	6	46	-	-	45	44	2
					35	32	4
					30	30	4
6	8	68	-	-	60	58	3
					50	46	4
					40	36	4
7	10	92	-	-	80	76	2
					70	68	4
					60	66	4
8	20	168	-	-	150	146	5
					145	138	5
					140	136	6
9	40	330	-	-	320	320	2
					310	306	3
					300	286	4

and South Pacific. pages 681 – 686, 2013.

[12] T. C. Xu, P. Liljeberg, and H. Tenhunen, Optimal Number and Placement of Through Silicon Vias in 3D Network-on-Chip, in IEEE14th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS), 2011, pp. 105 –110.

[13] K. Manna, V. S. Sri Teja, S. Chattopadhyay and I. Sengupta, TSV Placement and Core Mapping for 3D Mesh Based Network-on-Chip Design Using Extended Kernighan-Lin Partitioning, 2015 IEEE Computer Society Annual Symposium on VLSI.

[14] L. Chang and S. K. Lim, A design tradeoff study with monolithic 3Dintegration, in International Symposium on Low Power Electronics and Design, 2012, pp. 1–6.

[15] C. H. Lee, T. Y. Huang, C. H. Cheng, S. H. Huang, A Post-Processing Approach to Minimize TSV Number for High-Level Synthesis of 3D ICs. 2010 International Symposium on Computer, Communication, Control and Automation.

[16] D. Abts, N. D. E. Jerger, J. Kim, D. Gibson, and M. H. Lipasti, Achieving predictable performance through better memory controller placement in many-core cmps, in Proceedings of the 36th annual international symposium on Computer architecture, June 2009, pp. 451–461.