Designing 3D tree-based FPGA: Interconnect Optimization and Thermal Analysis

Vinod Pangracious, Habib Mehrez, Zied Marrakchi

HAL Id: hal-00873274
https://hal.archives-ouvertes.fr/hal-00873274

Submitted on 15 Oct 2013

Designing 3D Tree-based FPGA: Interconnect Optimization And Thermal Analysis

Vinod Pangracious and Habib Mehrez
LIP6, University of Pierre and Marie Curie Paris, France 75006
Email: Vinod.Pangracious@etu.upmc.fr

Zied Marakchi
FlexRas Technologies Paris, France
Email: Zied.Marrakchi@flexras.com

Abstract—The CMOS technology scaling has greatly improved the overall performance and density of Field Programmable Gate Array (FPGA), nonetheless the performance gap between FPGA and ASIC has remain very wide mainly due to the programmable interconnect overhead. Three-Dimensional (3D) integration is a promising technology to reduce wire lengths. We present a 3D design optimization methodology leveraged on Through Silicon Via (TSVs) to re-distribute the Tree interconnects into multiple stacked active layers using a tree-level horizontal break-point based on interconnect delay and to optimize the inter-layer heat dissipation. Nonetheless TSVs require a significant silicon area compared to planar interconnects and also bring critical challenges to the design of 3D ICs. In this paper we propose an architectural level interconnect and area optimization solution to minimize TSV count and programmable interconnects without compromising the FPGA performance. TSVs are also used very effectively to control the increase in inter-layer temperature of 3D ICs. We propose a TSV based 3D thermal optimization model for Tree-based FPGA. The experimental results from 3D Tree-based FPGA shows a 40% reduction of TSV count, 37% reduction in interconnect area and 28% reduction in power consumption.

I. Introduction

Three dimensional integrated circuit (3D-IC) technology has emerged as one of the most promising solutions for overcoming the challenges in interconnection and integration complexity in modern circuit designs [3]. 3D technology can effectively reduce global interconnect length and increase circuit performance. TSV is the key enabling technology for 3D integration, which is currently being actively evaluated as a potential solution to reduce the interconnect delay and increase the logic density in FPGA. Recent studies show that in Mesh-based FPGAs, 60% of overall design delay and 90% of the chip area are attributed to programmable routing resources [1]. It has also been reported that in Mesh-based FPGA, as much as 80% of the total power consumption is associated with programmable routing resources. Considering the area, delay and power consumption overhead, the programmable routing resources constitute the key design element in FPGA design.

Considering the programmable interconnect overhead, FPGA is an ideal device that can benefit significantly by 3D integration, in which the circuits are integrated vertically by stacking multiple dies together [3]. Since FPGA is an interconnect dominated device, it is essential to minimize the TSV count because TSV is much larger in size compared to silicon via and it also reduces flexibility in design, placement and routing during physical synthesis [4]. Moreover higher TSV count will have a negative impact on die yield of 3D manufacturing process as well as area of the chip. In this paper
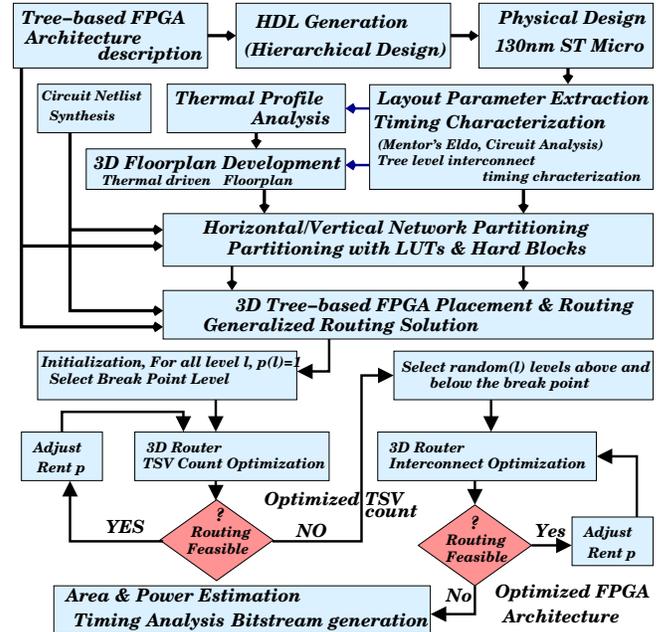

Fig. 1. 3D Tree-based FPGA design optimization methodology

we propose an innovative methodology to minimize TSV and local interconnect count of 3D Tree-based FPGA architecture.

II. Design And Optimization Methodology

A programmable Tree-based FPGA architecture is proposed in [6], where logic blocks and routing resources are sparsely partitioned into a multilevel clustered structure. The proposed methodology for design and optimization of 3D Tree-based FPGA architecture is illustrated in Figure 1. The HDL generator is designed to generate VHDL code based on a hierarchical design approach that partitions the design into smaller sections, implements them separately and assembles them together at the final design phase. The physical design experiments are performed based on the layout generated using Global Foundries 130nm technology node. Mentor's circuit simulator *Eldo* is used to estimate the wire delay and power consumption of switches and interconnection networks at different tree levels. The thermal model presented in [10] is used to extract the thermal profile of the multi-layer chip based on geometrical layout features and power consumption of the functional unites. The goal is to re-distribute the tree interconnect into two active layers in order to minimize the
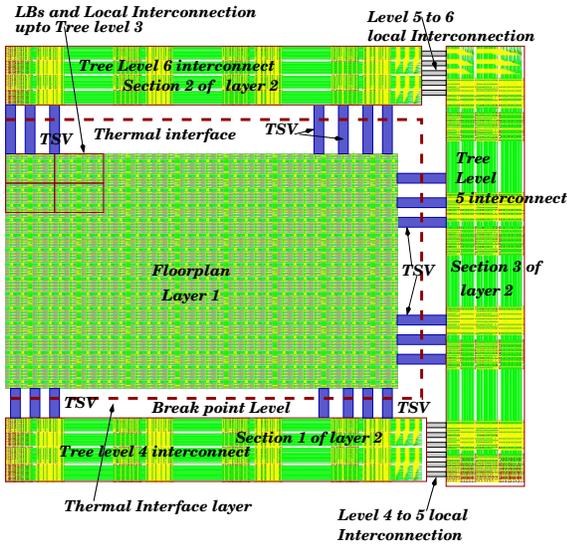
Fig. 2. 3D Tree-based FPGA with break-point between *level 3 and 4*

Fig. 3. 2 layer 3D Tree-based FPGA with interconnect levels 4 5 and 6 place in layer 1

interconnect delay and balance the temperature uniformly across the active layers of the 3D stacked Tree-based FPGA. The 2D design is partitioned at a particular level called the *break point level*, and interconnected using TSVs to optimize the delay of break point level and above. To optimize the network delay, the 2D design is partitioned between *levels 3 and 4* at which the network delay measured is more than *2ns* to form a 2 layer 3D stacked Tree-based FPGA. To illustrate the tree interconnect re-distribution process, a *7*-level Tree-based FPGA architecture is presented in Figure 2, where the break point is shown in dashed lines between *levels 3 and 4*.

To generate the 2 layer 3D Tree-based FPGA floorplan shown in figure 2, we used a thermal driven floorplanning tool [5] configured with Global Foundries 130*nm* technology node. This tool is configured to optimize wire length and temperature of the block level schematic of the 3D stacked Tree-based FPGA chip. The floorplan tool takes account of allowable aspect ratios, connectivity between blocks and power consumption of each functional block for optimization. In this study two floorplans were generated and illustrated in Figure 3. The first floorplan consists of the logic units and local interconnections up to *level 3*. The second floor plan consists of interconnect *tree levels 4 and above*. The floorplan tool generates the thermal profile and the wire delay of local and global interconnect metal layers. A 3D programmable interconnection network for Tree-based FPGA architecture is presented in [5]. As illustrated in Figure 3, the programmable interconnects of a Tree-based FPGA are arranged in a multi-level network with the switch blocks placed at different tree levels using a *Butterfly-Fat-Tree* network topology. The 3D interconnect network organization and design validation are presented in [5]. The inter-layer vias (TSV) are to be limited because they are large in size compared to the minimum feature size on the die, while the finest vias currently available are about 1μm X 1μm with a pitch of about 1.6 to 3μm [2].

Although the design engineers are trying to reduce the TSV dimensions, the minimum feature size on the die is also shrinking. Therefore the TSVs are expected to remain larger than the wire dimensions in metal layers within the die. This
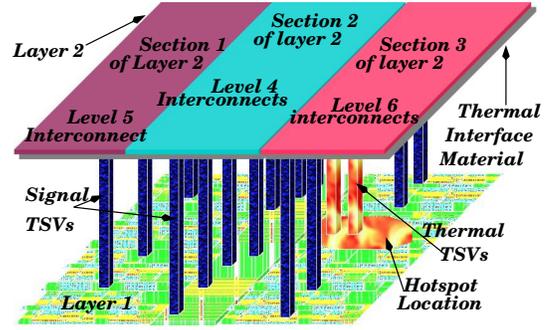
means it is essential to minimize the TSV count to reduce design complexity and cost of manufacturing. In [8] a 3D place and route (TPR) is presented to investigate the wire length and delay associated with 3D Mesh-based FPGA. TPR [8] assumes the number of TSVs is the electrical equivalent of the horizontal channel width. In 3D integration technology, the TSVs are much thicker than horizontal wires, which makes this assumption impractical. A design framework for 3D FPGA architecture level exploration methodology is presented in [7]. It includes an additional feature to explore vertical interconnect distribution, however this leads to usage of 2D and 3D switch blocks intermittently, which may lead to number of design and manufacturing issues. Thermal issues in FPGAs have been relatively unexplored for a long time. One report [9] proposes the use of distributed sensors for monitoring temperatures in FPGAs. However, it considers only configurable logic blocks (CLBs) in the FPGA fabric, and consequently, observe very little temperature variations across the die. Nevertheless the programmable interconnect network of FPGA consumes a lot of power. In contrast, the methodology we use considers the power consumption of computational and interconnect sections separately to investigate the temperature variations in FPGA. First we measure the 2D block level temperature estimation and in a next step we use the proposed special TSV placement and positioning algorithm to evaluate the temperature variation of the 2-layer 3D stacked Tree-based FPGA system.

III. TSV COUNT OPTIMIZATION

Since FPGA is an interconnect dominated device, it is essential to minimize the TSV count because TSVs consume more silicon area than horizontal interconnects. The vertical and horizontal interconnect optimization is performed using Rent's parameter [6] "p" defined for a Tree-based architecture shown in equation 1. The Tree level is represented as ℓ and k is the cluster arity, c is the number of in/out pins of an LB and IO is the number of in/out pins of a cluster located at level ℓ.

$$IO = c.k^{\ell \cdot p} \tag{1}$$

In a Tree-based FPGA architecture [6], The LBs (Logic Blocks) are grouped into clusters located at different levels. Each cluster contains a switch block to connect local LBs. A switch block is divided into MSBs (Mini switch Blocks). The Tree-based FPGA architecture unifies two unidirectional upward and downward interconnection networks using a *Butterfly-Fat-Tree* topology to connect DMSBs (Downward

MSBs) and UMSBs (Upward MSBs) to LB inputs and outputs. For a downward MSB (DMSB) interconnection network, a reduction in the number of inputs at level ℓ impacts level $\ell + 1$, since the number of DMSBs at level $\ell + 1$ is equal to the number of inputs at level ℓ. In the case of the upward interconnection network, a reduction in number of outputs at level ℓ impacts at level $\ell + 1$ since the number of UMSBs at level $\ell + 1$ is equal to the number of outputs at level ℓ. The Rent's parameter [6] based upward and downward network model is explained below.

A. The Interconnect Network Model

A cluster situated at level ℓ contains $N_{in}(\ell - 1)$ DMSB, where $N_{in}(\ell)$ is the number of inputs of cluster located at level ℓ with k outputs and $\frac{N_{in}(\ell) + kN_{out}(\ell-1)}{N_{in}(\ell-1)}$ inputs, whereas k is the cluster arity size. Since DMSBs are full crossbar devices, the total number of switches at level ℓ cluster is $k(N_{in}(\ell) + kN_{out}(\ell-1))$. At each level ℓ, there are $\frac{N}{k^\ell}$ clusters, whereas N is total number of Logic Blocks and the total number of switches or interconnects in the downward network is

$$\sum_{\ell=1}^{\log_k(N)} k \times N \times \frac{N_{in}(\ell) + kN_{out}(\ell-1)}{k^\ell} \quad (2)$$

Following equation 1, we can simplify the number of outputs of a Logic Block $N_{out}(0) = c_{out}$ and the number of inputs equals $N_{in}(\ell) = c_{in}.k^{\ell.p}$ and $N_{in}(\ell-1) = c_{out}.k^{(\ell-1)p}$ and so on. The total interconnects used at each level ℓ can be calculated by equation 3.

$$N_{switch}(Tree) = N \times (k^p c_{in} + 2kc_{out}) \times \sum_{\ell=1}^{\log_k(N)} k^{(p-1)(\ell-1)} \quad (3)$$

The TSV count minimization methodology is developed using Rent's parameter-based iterative 3D router [5] program. The aim is find the best tradeoff between device routability and interconnect requirements of each MCNC application. The 3D TSV optimizer program as illustrated in Figure 1, first selects the horizontal break point level of tree interconnect to optimize the number of TSVs required between layer 1 and 2 of the 3D stacked Tree-based FPGA. After finishing the break point level, the interconnect optimizer chooses other levels above or below the break point in a random order, which could be either in active layer one or two to optimize the number of interconnect switches in the upward and downward interconnection network. The interconnect optimizer will consider the same architecture with different rent parameter p values for each iteration to find the optimum FPGA hardware required to implement each application netlist. Tables I and II present the TSV and interconnect optimization results. An average reduction of 40.1% TSVs and 37% reduction in interconnect area is recorded in our experimental results. The performance degradation measured for 3D Tree-based FPGA is 4.7% compared to 8.1% of the 3D Mesh-based FPGA [7] of same density. The experimental results confirm that 3D Tree-based FPGA is a consistent architecture in terms of performance and area to build high density and high performance FPGA, which is unlikely to be attained using Mesh-based FPGA architecture.

TABLE I.   3D TREE-BASED FPGA WITH 7 LEVEL AND ARITY 4: TSV COUNT OPTIMIZATION RESULTS

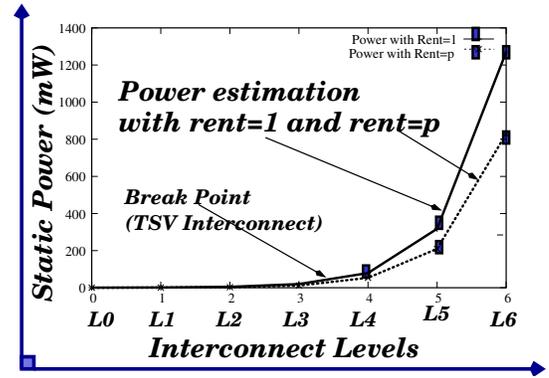| | Tree Levels=7, Arity=4, Arch=4x4x4x4x4x4x4 | | | |
	Tree-based FPGA		Tree	Mesh
Circuits	Optimized	$3D_{TSV}$	Speed	40% TSV Reduction
MCNC	Rent's "p"	Gain(%)	degradation%	Speed degrade%
alu4	0.47	53	4.5	2.8
apex2	0.51	49	6.2	5.4
apex4	0.61	39	1.1	5.9
bigkey	0.60	40	3.5	3.5
clma	0.58	42	5.1	9.6
des	0.56	44	4.1	9.5
diffeq	0.66	34	4.5	0.2
dsip	0.65	35	4.1	8.2
elliptic	0.66	34	3.4	9.6
ex1010	0.55	45	3.5	2.1
ex5p	0.58	42	5.1	0.4
frisc	0.62	38	5.4	4.5
misex3	0.64	36	5.8	15.1
pdc	0.59	41	3.8	1.8
s298	0.55	45	6.8	5.8
s38417	0.65	35	5.5	16.7
s38584	0.62	38	4.5	8.6
seq	0.61	39	5.5	5.9
spla	0.58	42	5.2	2.5
tseng	0.65	35	5.5	1.3
ava	0.63	37	5.3	8.8
average	0.598	40.1	4.7	8.1

Fig. 4.   3D Tree-based FPGA power estimation: Power consumption at different level with Rent=1 and Rent=p

TABLE II.   7 LEVEL TREE FPGA LAYOUT WITH 16K CLBS

Tree Levels=7, Arity=4, Arch=4x4x4x4x4x4x4				
Tree-based FPGA Block Type	3D Chip # Layer	Rent ="p" Avg(21 MCNC)	Block Area μm^2	Power mW
CLB Unit	1		93635273	0.15
Switch box Level 0	1	0.67	2412	0.25
Switch box Level 1	1	0.54	10800	0.83
Switch box Level 2	1	0.66	37496	2.70
Switch box Level 3	1	0.59	232128	13.13
Break-point	-	-	-	-
Switch box Level 4	2	0.67	6072770	47.46
Switch box Level 5	2	0.66	41553499	214.05
Switch box Level 6	2	0.65	42139683	838.97
Average		0.63		
Average Interconnect Power Reduction is 28.43%				

IV.   POWER OPTIMIZATION OF 3D TREE-BASED FPGA

The rent-based architecture optimization shows a reduction of 28.43% in total power consumption of the 7-level Tree-based 3D interconnect network. The TSV count at the break point level is reduced to 59.8% from the rent equal to 1 (100% TSV). The layer 1 and 2 optimization results are presented Table II. A reduction of 28.43% in total power
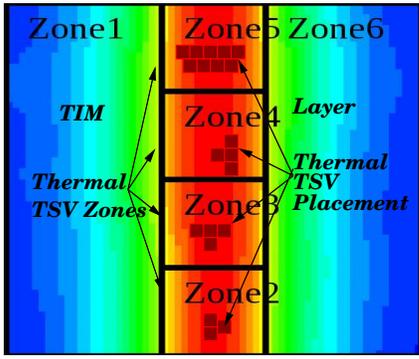
Fig. 5. Special TSV placement Zones at epoxy layer (Re-distribution Layer of Layer 1)

Fig. 6. 3D Thermal analysis of Tree-based interconnect levels and Logic units

consumption is achieved by reducing 37% of the interconnects and corresponding switches through an architecture optimization process. This is very promising for FPGA architecture, since FPGA is an interconnect-dominated architecture and it is impossible to manufacture it with huge numbers of TSV and switches. Figure 4 presents the analysis of power estimation with 100% and 59% TSVs. As the tree grows to higher levels, the number of interconnects and switches also grows exponentially. Table II shows the tree level power distribution.

V. TSV Based FPGA Thermal Control

The thermal model used in this work is presented in [5] and [10]. The 3D thermal model considers the effective thermal conductivity of each tier in the 3D stack and also the flexibility to include special TSV zones [10] for temperature control as illustrated in Figure 5. The special TSV placement is done in accordance with the TSV count optimization result. The 3D FPGA thermal model has the flexibility to decide the location based on the thermal profile of the chip to re-distribute inter-layer temperature. In certain locations the signal TSVs helps to optimize the inter-layer temperature, however the 3D thermal model uses a limited number of thermal TSVs in a few localized spots where no signal TSVs exist. However this is a limited operation, as TSVs are expensive and cause degradation in the performance of active devices. The effective thermal conductivity of active and support layers in the 3D stack are computed by using equation 4. TSV density is computed based on TSV count, dimension, and pitch constraints [11] between TSVs of layer 1 and 2. The k_{cu} and K_{th} are the thermal conductivity of copper TSVs and support or glue material.

$$k_{eff} = k_{cu} \times (TSV_{Area}) + K_{th} \times (Level_{BPArea} - TSV_{Area})$$
(4)

In our 3D thermal experiments, we considered all FPGA blocks like BLE and interconnect levels ranging from 0 to 6 in the tree structure. Table II presents the optimized area and power of each block and which layer it is placed at in the 3D stacked Tree-based FPGA. Layer 1 consists of BLEs and local interconnects up to level 3 and layer 2 consists of higher level interconnection network of the Tree architecture. The thermal analysis structure is configured using face-to-back 3D integration method. The peak temperature increases from 345K to 356K for the 2 layer 3D stacked Tree-based FPGA and average temperature is 348K. The temperature analysis of the two layer 3D Tree-based FPGA is presented in Figure 6.
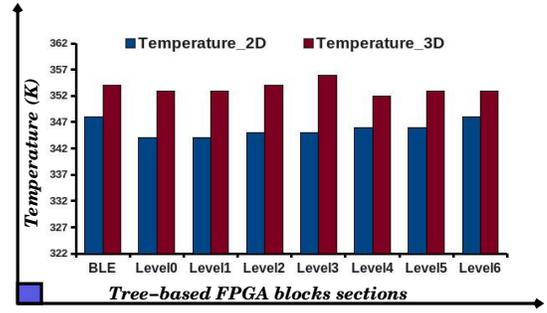
With our localized TSV placement and positioning algorithm, the peak and average temperature of 3D stacked Tree-based FPGA are maintained at 356K and 348K respectively.

VI. Conclusion

An architecture level design and optimization 3D Tree-based FPGA and thermal analysis were presented. The issues associated with TSV size and count and its impact on design of 3D integrated circuits were studied and presented. The study reveals TSV management and thermal control of 3D stacked chip is essential for guaranteed performance and yield. The design methodology adopted based on Rent-parameter shows a performance degradation of 4.7% for an average reduction of 40.1% TSVs. Furthermore the 3D thermal model is used in accordance with TSV reduction to optimize the inter-layer temperature. These results make 3D Tree-based FPGA as a viable alternative to build 3D re-configurable systems.

References

[1] A DeHon, *Reconfigurable Architectures for General-Purpose Computing*, Ph.D dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1996.

[2] ITRS, *International technology roadmap for semiconductors*, 2012 [Online]. Available: http://public.itrs.net

[3] A Rahman, R Reif, *System Level Performance Evaluation of Three Dimensional Integrated Circuits*, IEEE Trans VLSI System, Vol 8, PP 671-678, Dec 2000.

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

[5] V Pangracious, Z Marrakchi, E Amouri, and H Meherez, *Performance Analysis and Optimization of High Density Tree-based 3D Multilevel FPGA*, ARC 2013, pp 197-209, 2013.

[6] Z Marrakchi, H Mrabet, U Farooq, and H Mehrez: *FPGA Interconnect Topologies Exploration.* Int. J. Reconfig.Comp.2009

[7] K Siozios, A Bartzas, D Soudris. *Architecture Level Exploration of Alternative schmes Targeting 3D FPGAs: A Software Supported Methodology*, International Journal of Reconfigurable Computing,Vol2008.

[8] C Ababei, Y Feng, B Goplen.*Placement and routing in 3D integrated circuits*, IEEE Design & Test of Computers, vol.22, no.6, pp.520531,2005.

[9] S Velusamy et al. *Monitoring temperature in FPGA based SoCs*, ICCAD San Jose, CA, 2005.

[10] J Ayala, A Sridhar, V Pangracious, D Atienza, Y Leblebici, *Through Silicon Via-Based Grid for Thermal Control in 3D Chips* NanoNet pp.90-98 2009.

[11] D M Jang, C Ryu, K Y Lee, *Development and evaluation of 3-D SiP with vertically interconnected Through Silicon Vias (TSV)* Proceedings of the 57th ECTC '07, pp.847-852, USA, May-June 2007.