

Thermal-aware partitioning for 3D FPGAs

Chaitanya Nunna, Krishna
Department of Advanced Informatics, Kyushu University

Mehdipour, Farhad
E-JUST Center, Kyushu University

Murakami, Kazuaki
Department of Advanced Informatics, Kyushu University

<https://hdl.handle.net/2324/26567>

出版情報 : Field Programmable Logic and Applications (FPL), 2012 22nd International Conference on, pp.475-476, 2012-08. Institute of Electrical and Electronics Engineers

バージョン :

権利関係 : (C) 2012 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.



THERMAL-AWARE PARTITIONING FOR 3D FPGAs

Krishna Chaitanya Nunna[†], Farhad Mehdipour^{*} and Kazuaki Murakami[†]

[†] Department of Advanced Informatics, Kyushu University, JAPAN

Email: krishna@soc.ait.kyushu-u.ac.jp

^{*} E-JUST Center, Kyushu University, JAPAN

ABSTRACT

Three-dimensional FPGA is one of the promising innovations which can lead to the reduction in delay, area and power. There is an absolute necessity to develop algorithms and software tools to exploit the advantages of the third dimension, and to solve complex tasks associated with them. Also, thermal issues are cited as critical concern in 3D integration which results in degradation of device performance. In this paper we are proposing an idea for thermal-aware partitioning targeting the power/thermal-aware EDA flow for 3D FPGAs.

1. INTRODUCTION

The main focus of commercial vendors and also the majority of published research on FPGA architecture and EDA are improving area efficiency and performance. Though area and speed have been the main research focus to date, power is likely to be a key consideration in the design of future FPGAs especially in three-dimensional platform. In spite of advantages over 2D, thermal effects are expected to be significantly higher in 3D chips due to the higher power density and this can cause greater degradation in device performance and chip reliability. Additionally, excessive heat dissipation is one of the most critical challenges today. Concentration on power and thermal issues at the early stage of design cycle will reduce the burden on later stages which are already having bigger problem size. Through our research, we are proposing an EDA methodology for 3D FPGA for handling the thermal effects at the partitioning, placement and routing stages depending on the power density values. In a typical FPGA design flow, the goal of the placement is to distribute the design among the physical logic blocks in the reconfigurable fabric and to minimize the delay along the critical path and enhance resulting circuit routability. Partitioning of the design happens implicitly with placement. In 3D case, explicit partitioning of design (Fig. 1) gives the opportunity for reducing the workload on placement by handling typical issues at the earliest.

The past and recent work related to this area is encouraging. Ababei *et al.* [1] introduced 3D place and route tool for FPGAs in which the partitioning is done explicitly. In [2] a new temperature-aware placement and routing algorithm targeting 3D FPGAs has been proposed which also depicts the explicit partitioning. Only few

papers are stated here due to space limitation. Even though these researchers considered explicit partitioning, none of them dealt with critical issues related to the power and the temperature.

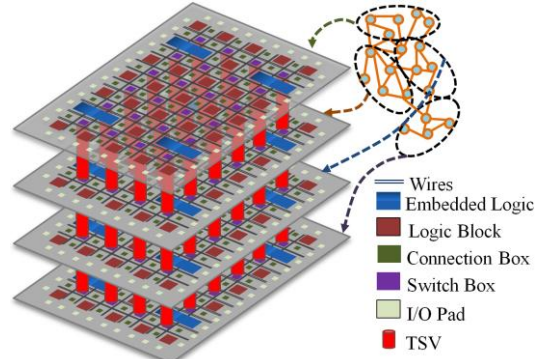


Fig. 1 (a) Explicit partitioning and layer assignment

For our target 3D EDA flow, the architecture is similar to Xilinx Virtex-II device, which is a commercial 2D FPGA, for each layer and such identical device layers are aligned vertically using through-silicon-vias (TSVs). As the TSVs handle critical signal transition between layers, their effect on thermal issues should also be considered. In order to add the effect of TSV utilization, switching elements associated with them need to be known. The proposed EDA methodology in section 2 handles these issues in calculating the power densities.

2. THERMAL-AWARE PARTITIONING

Due to remarkable complexity of recent designs, it seems necessary to estimate performance metrics in early design phases. The availability of accurate early estimates can reduce the required number of iterations through the entire design flow, which allows design trade-offs to be evaluated at an abstract level resulting in reduction in the design time and cost.

As shown in Fig. 2(a), we are targeting to develop a thermal-aware 3D FPGA design flow based on early estimations of power and thermal values. Proposed flow follows traditional 2D design flow except at the stages which need architectural changes. To solve the relating issues, we are aiming for efficient power/thermal model at partitioning, placement and routing stages. Partitioning is followed by layer assignment which will distribute the design partitions among layers depending on the thermal

values. Power/thermal models will be updated for improved accuracy throughout the design flow. Our methodology is based on TPR [1] in which the existing hMeTiS based partitioning will be replaced by the modified version of the recursive partitioning algorithm. The implementation flow for thermal-aware partitioning is shown in Fig. 2(b).

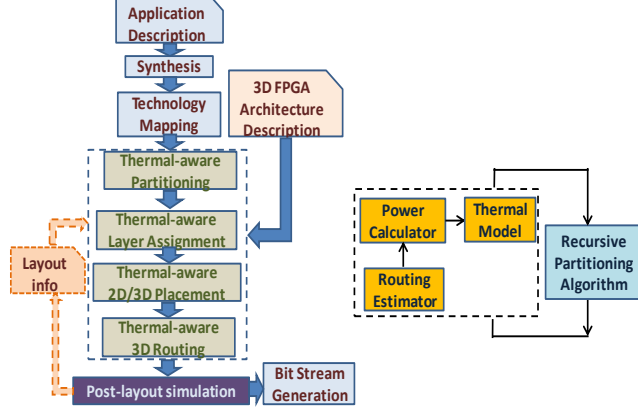


Fig. 2 Proposed (a) thermal-aware 3D FPGA design flow and (b) thermal-aware partitioning methodology

On an FPGA, temperature can be analyzed as follows:

$$T = \sum_{i=1}^N P_{D,i} * R_i \quad (1)$$

In (1), T is the chip temperature, N is the number of layers (number of partitions is equal to number of layers in 3D case), $P_{D,i}$ is the power density and R_i is the thermal resistance of each layer. Power is again divided into static and dynamic components. From our experiments on power consumed by different resources of FPGA, it is observed that the dynamic power due to the routing resources is dominating compared to any other resources and this is increasing with the increase in circuits size (Fig. 3).

Dynamic power depends mainly on the switching activity of the signals (α) and switching capacitance (C) as defined below:

$$P_{dynamic} = \alpha * f * C * V^2 \quad (2)$$

Where f is the frequency of operation and V is the supply voltage and are constants for a given architecture. At the partitioning stage, since no geometrical information of cells is available, there should be proper estimation criteria for switching capacitance C . For this, we are developing a methodology for early estimation of hierarchical routing resources. From our observations on these resources with respect to the power estimations, the dynamic power for the routing resources is modeled as:

$$P_{routing} = N_{sg}P_{sg} + N_{swt}P_{swt} + N_{cut}P_{(sw+TSV)} \quad (3)$$

Here N_{sg} is the number of routing segments used by the given circuit, P_{sg} is power associated with each segment (metal wire) and it is technology dependent. Depending on N_{sg} , associated switches N_{swt} can be estimated and so the respective power P_{swt} . N_{cut} is the cut size (number of nets

involved in intra-layer communication) which can be obtained from the partitioning of the circuit and $P_{(sw+TSV)}$ is the power of the switching element in presence of TSV. For estimating N_{sg} , we are extending the work proposed by [3] separately for 2D and 3D nets.

3. PRELIMINARY RESULTS

To get the knowledge on power distribution among FPGA resources, we have conducted experiments on five MCNC benchmark circuits using VPR with the power model proposed by Poon *et al.* [4]. Our results are shown in Fig. 3. We can see that the power consumed by routing resources is much higher than the other and it is growing with the increase in circuit size. On the other side, power due to logic is low and almost unchanged. From this we have observed that the role of routing fabric in FPGA power is momentous.

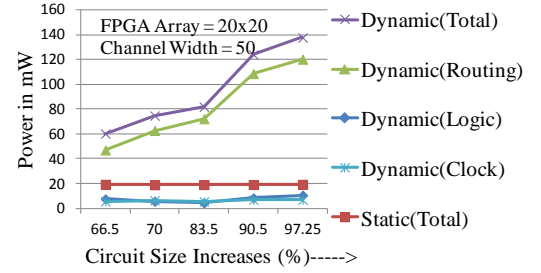


Fig. 3 Power components in FPGA

4. CONCLUSION AND FUTURE WORK

The implementation of the proposed work is under development. At first we are aiming to complete the thermal-aware partitioning depending on the routing estimations. We have proposed our methodology for early estimation of hierarchical routing resources and submitted a paper on the same. In near future we will use the same model (updated accordingly) for the placement phase and will try to improve the accuracy in estimating the required parameters.

5. REFERENCES

- [1] Ababei. C, *et al.*, "Three-dimensional place and route for FPGAs", *IEEE Trans. CAD of Integrated Circuits and Systems*, pp. 1132 – 1140, June 2006.
- [2] Kostas. S and Dimitrios. S, "A Temperature-Aware Placement and Routing Algorithm Targeting 3D FPGAs", in *Proc. 16th IFIP WG 10. 5/IEEE Int. Conf. VLSI-SoC*, 2008, pp. 211–231.
- [3] Dai Zhibin and D.K. Banerji, "Routability Prediction for Field-Programmable Gate Arrays with a Routing Hierarchy", in *Proc. 16th IEEE Int. Conf. VLSI Design*, Jan. 2003, pp. 85-90.
- [4] K. Poon, A. Yan and S.J.E. Wilton, "A Flexible Power Model for FPGAs", in *12th Int. Conf. Field-Programmable Logic and Applications*, Sept 2002, pp. 312-321.