SAME 2008 Forum

University Booth

Interconnect Explorer: a High-Level Estimation Tool for On-Chip Interconnects

Antoine Courtay[1,2], Olivier Sentieys[2], Johann Laurent[1] & Nathalie Julien[1]
[1]Université Européenne de Bretagne – UBS – Lab-STICC – Lorient, France
[2]University of Rennes – IRISA – Lannion, France

Abstract

It is now well admitted that interconnects introduce delays and consume power and chip resources. To deal with these problems, some studies have been done on performance optimization. However, such techniques are not always based on good criteria for interconnect performance optimizations. We have, therefore, developed a high-level estimation tool based on transistor-level characteristics, which provides fast and accurate figures for both time and power consumption. These results allowed us to determine new key issues that have to be taken into account for future performance optimizations.

Table of Contents of the paper

About the Authors

Antoine Courtay received B.S. and M.S. degrees in Electronics from the University of South Brittany, Lorient, France. He is currently working toward the Ph.D. degree at the Lab-STICC (Laboratoire des Sciences et Techniques de l'Information, de la Communication et de la Connaissance), Université Européenne de Bretagne – UBS, Lorient and at the IRISA (Institut de Recherche en Informatique et Systèmes Aléatoires), University of Rennes, Lannion. His research interests include power/energy modelling and optimization of on-chip buses.

1. Introduction

Today, System on Chip (SOC) are more and more complex and require many computational resources, implying a large volume of data to be stored or to be transmitted. To transfer this data from memory to processor or from one processor to another, on chip interconnect buses or networks have to be used. In state-of-the-art SOC, interconnect can represent up to 50% of the total power consumption [1,2]. Moreover, the transistor and wire dimension scaling has a strong impact on propagation time; indeed the propagation time of a wire becomes higher than that of a gate [2]. Therefore, power and delay estimation and optimization due to interconnections has become a major issue in SOC design. With the increase of the die size and the device count, more wires (which are getting longer) are needed for interconnections. It is thus essential to take interconnect power consumption and delay into account during the first design stages of a system.

In this paper, we propose, after the presentation of the power consumption characterization flow for buses, an estimation tool that allows the user to obtain numerous results about the power consumption of the interconnect networks. Then, based on the analysis of the results provided by the tool, we suggest new ways to optimize the interconnect performance (delay and power consumption).

2. Power consumption characterization

The first step for interconnect modelling is to represent the interconnect behaviour as realistically as possible. In order to obtain the highest precision in time and power consumption, experiments must be carried out at the physical level. Therefore we decided to model interconnects at the transistor level using a SPICE simulator.

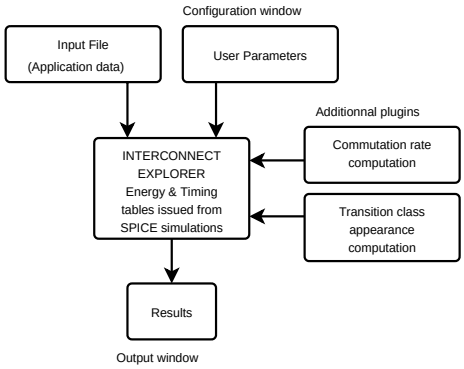
Physical parameters which allow wires to be modelled are:

- *R*, wire resistance, expressed in Ohm [Ω];
- *C*, wire capacitance, expressed in Farad [F];
- *L*, wire inductance, expressed in Henry [H].

Their variation depends on the wire characteristics (metal composition, wiring level) as well as dimensions. Inductance has an impact only for very deep submicron technologies (below 45 nm) and for extremely long wires [3].
Furthermore, the simulations presented in this paper satisfy the conditions proposed in [4] that define the range of the interconnect length in which inductance effects are not significant. Thus, considering only an *RC* model for the wire gives accurate results for technologies and bus length used in this paper. It is possible to characterize the wire with elementary parameters which can be found in manufacturer's Design Kits.
The distribution of *R* and *C* on the wire in order to model its behaviour as accurately as possible must be considered. The lumped model is a simple interconnect model which consists in putting end-to-end the values of *R* and *C* found previously. However, its precision is much less reliable in terms of propagation time than a model, where *R* and *C* are distributed. We have retained the $\pi 3$ model (which consists in splitting the wire resistance into three and the wire capacitance into four) for our experiments because of its simplicity and its precision (estimation error of time less than 5%) [5].
An *n* bit bus is considered. It consists of *n* parallel wires of the same length and at least 2*n* buffers (*n* input buffers, *n* output buffers and others if bufferization is used) for data propagation between two cells. Using several wires gives rise to a new capacitive coupling between the wires known as the crosstalk capacitance. The effects due to crosstalk can be summarized into three categories:
• The first one is that crosstalk induces noise; the coupling capacitance between adjacent wires introduces a permanent link between them. When a transition occurs on a wire, its neighbours are affected, because a voltage peak is generated on them [6].
• A second issue is the increase in propagation time. When two wires are switched simultaneously, a voltage peak is generated. This peak can, according to the configuration of transitions, slightly accelerate (in the case of simultaneous transitions in the same direction) or slow down (in the case of simultaneous transitions in opposite side) the propagation on the victim wire
• Finally, the last issue is the increase in power consumption. Indeed, the power consumption depends linearly on the capacitance presented by a device. Since the wire capacitance depends on the crosstalk capacitance value, the crosstalk contributes to the increase in the dynamic power consumption [7].
Knowing the physical parameters of the wires and the bus, the propagation time and power consumption can be modelled. The first step of the modelling process is to identify which parameters impact significantly on delay and power consumption.
The first parameter is the technology used and its associated number of metal layers.
The second parameter is the wire length since this parameter impacts on capacitance and resistance.
Since crosstalk capacitances have effects on power consumption and propagation time according to the configuration of transitions, the different kinds of transitions are also parameters to be modelled
Using these parameters, power consumption and delay modelling can be realized at the circuit level using SPICE simulations (we used ELDO v5.7 in this paper). These simulations have been done for three different technologies (130, 90 and 65 *nm*). The results obtained with SPICE, in terms of time and energy consumption, have been summarized in multi-input tables for various previously mentioned parameters. These tables are used by the high-level estimation tool that will be presented in the next section.

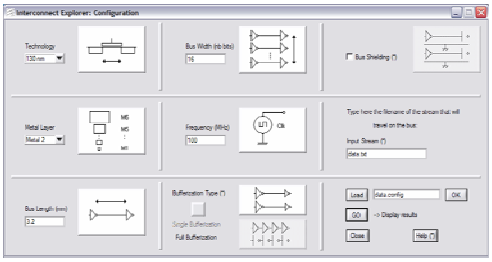
3. Power consumption estimation

A tool, called Interconnect Explorer, has been developed for high-level estimation of interconnect performances. This tool is based on energy and timing multi-input tables. These tables depend on input parameters (technology, metal layer, wire length, buffer and repeater size, transition type) and their values are obtained with transistor-level characterization. The estimation flow used by Interconnect Explorer is explained in the next figure and detailed below.

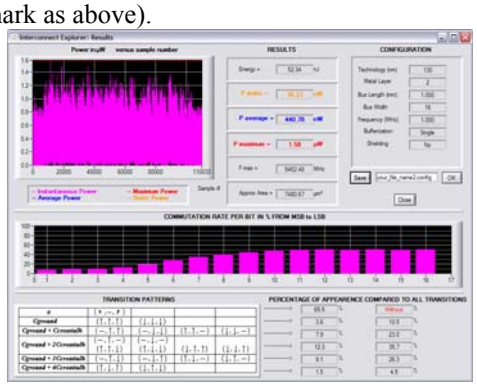


When using Interconnect Explorer, users have to choose their bus configuration by setting the following parameters in the tool configuration window (see next figure): technology, metal layer, bus length, bus width, frequency, and bufferization type. Users have also to provide an input file which contains the data that the bus is handling. Some additional plug-ins have been included in this tool to compute commutation rate per bit on the bus as well as the probability of the appearance of each transition class. Commutation rates per bit are obtained by using the data input file. We compute the activity on each wire from the ratio of the

number of transitions on the wire to the total number of data. Similarly, the probability of appearance of each transition class is obtained by computing the ratio of the number of occurrences of each transition class to the total number of transition occurrences.



After configuration, Interconnect Explorer provides users with, in the output window (see next figure), results in terms of energy consumption, static power consumption, average dynamic power consumption, maximum dynamic power consumption, instantaneous dynamic power consumption, maximum frequency allowed on the bus (determined by worst case transition), area on the bus (area for wires and buffers), commutation rate per bit (useful to evaluate performance optimization techniques), and percentage of appearance of each kind of transitions (the same remark as above).



The maximum error between consumption results provided by Interconnect Explorer and SPICE simulation is less than 6%. Interconnect Explorer provides results instantaneously (less than 1 second computation) whereas a SPICE simulation of the same configuration takes several hours.

4. Power consumption optimization

When using the Interconnect Explorer tool to analyse state-of-the-art performance optimization techniques some new issues for power consumption optimization have been underlined in [8]

• It can be noticed that transitions which are eliminated by performance optimization techniques are not appearing very often so, do not only focus on low arising consuming transitions since they are not predominant in the total transition number.

• Focus on the lines with the largest data activity because it can be noticed that applying performance optimization techniques on least significant bits has a better impact in terms of power consumption reduction.

• Avoid falling transitions as much as possible: a key point for power optimization can be to encode data such as falling transitions on the bus are achieved with the lowest crosstalk capacitance and thus consume less energy as possible.

• Design codec with power overhead as low as possible, and therefore focus on very simple techniques because most of the state-of-the-art techniques do not always take the extra power consumption due to codec into account or are not efficient for reasonable bus length [9].

Based on these issues extracted from [8], a novel optimization technique has been proposed in [10].

References

[1] N. Magen, A. Kolodny, U. Weiser and N. Shamir, Interconnect-power dissipation in a microprocessor. *Proceedings of the international workshop on System level interconnect prediction* (2004), pp. 7-13.

[2] R. Ho, K. Mai and M. Horowitz, The Future of Wires. *Proceedings of the IEEE* (2001), Vol 89, pp. 490-504.

[3] W.J.Dally and J.W .Poulton, Digital Systems Engineering, Cambridge University Press (1998).

[4] Y.I.Ismail, E.G.Friedman, and J.L.Neves, Figures of merit to characterize the importance of on-chip inductance. *IEEE Trans. On VLSI* Vol 7, No 4, pp.442-449 (1999).

[5] J.M.Rabaey, A.Chandrakasan, and B.Nikolic, Digital Integrated Circuits: A Design Perspective, Pearson Education (2003).

[6] A.Devgan, Efficient coupled noise estimation for on-chip interconnects, *Proceedings of the IEEE/ACM International Conference on Computer Aided Design* (1997), pp.147–153.

[7] C.Duan and S.P .Khatri, Exploiting crosstalk to speed up on-chip buses, *Proceedings of the Conference on Design, Automation and Test in Europe* (2004), Vol.2, pp.778–783.

[8] A. Courtay, O. Sentieys, J. Laurent and N. Julien, High-Level Interconnect Delay and Power Estimation, *J. Low Power Electronics* (2008), Vol. 4, No. 1, pp.21–33

[9] C. Kretzschmar, A.K. Nieuwland and D. Muller, Why transition coding for power minimization of on-chip buses does not work, *Proceedings of the Conference on Design, Automation and Test in Europe* (2004), pp.10512–10517

[10] A. Courtay, J. Laurent, O. Sentieys and N. Julien, Novel Cross-Transition Elimination Technique Improving Delay and Power Consumption for On-Chip Buses, *accepted for the International Workshop on Power and Timing Modeling Optimization and Simulation* (PATMOS)