1

Review of the literature

Part I: Microelectronics design

Chapter abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

1.1 Introduction

1.2 2D architecture and its limitations

In order to continuously improve the performance of integrated circuits (IC), technologists deploy enormous efforts to produce IC manufacturing process that is compelling to follow the well-known Moore's Law (see Fig. 1.1). This empirical

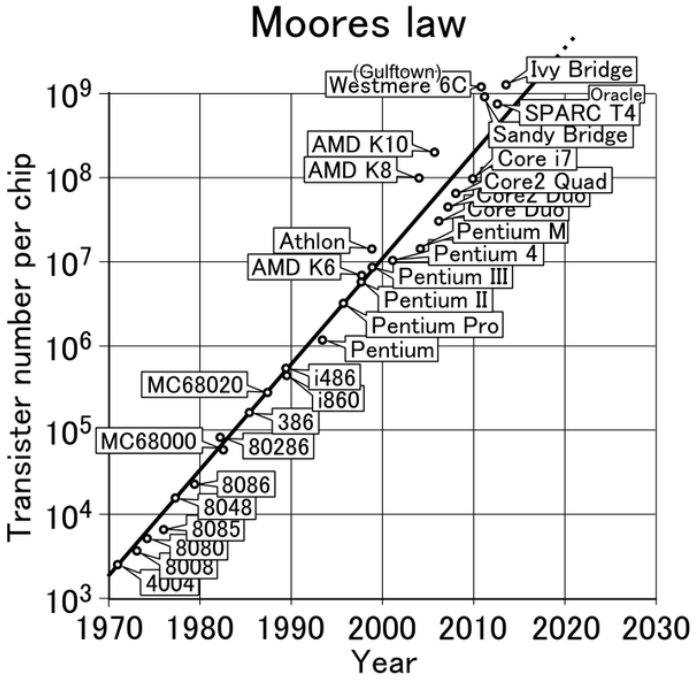

Figure 1.1: Moore's law [1]

law predicts a doubling of the transistors' integration each 18 months and therefore increasing logic capacity of the circuit per unit area.

The improvements of 2D architectures are primarily driven by the reduction of the transistor size. By reducing transistor dimensions, the switching speed is increased thanks to the shorter distance between the source and the drain, implying an improvement of the overall speed of the designs.

However, as the transistor size is decreasing, the observed improvement is also getting smaller. Indeed, a smaller transistor allows higher device density but will slightly increase the total delay (sum of gate and interconnection delays). Also, power consumption is also increased due to higher leakage and increasing interconnection wire length [2].

With the miniaturization, quantum effects such as quantum tunnelling will significantly affect how a transistor behave [3].

In addition to these physical aspects, economical considerations that will hinder the IC evolution beyond 20nm have to be taken into account [2, 4].

In order to overcome these limitations, new technologies have been proposed such as the carbon nanotubes [5], the nanowire transistors [6], the single-electron transistors [7], but also the 3D-Stacked Integrated Circuits (3D-SIC) proposed by the

academic and industrial communities. The latter has been often cited as the most prominent one [8].

1.3 3D integration

Most of the current ICs are designed with electronic components (i.e. transistors) that are planar (although multi-gate transistors, such as finFETs tends to extend in the 3rd dimension) interconnected using up to a maximum of 12 (also planar) wiring (metal) layers per circuit. Those conventional ICs can thus be considered to be two-dimensional (2D)-ICs since the interconnections are predominantly made in a planar fashion [9, 10]. As a major evolution of 2D-ICs, 3D-SICs are designed with multiple traditional 2D-ICs (that are manufactured independently, using standard CMOS technology) that are assembled (stacked) vertically in 3D-tiers. Different 2D circuits communicate between tiers using vertical interconnections that need to connect front side of the chip and the backside, i.e. they need to traverse bulk silicon. These connections, known as – Through Silicon Vias (TSV), can be today manufactured with satisfactory geometrical properties, namely their diameter, pitch and height, allowing efficient integration of real-world systems [11, 12]. This is shown in Fig. 1.2, where 2 dies, oriented face down are connected. An active component (i.e. logic gate) of the T1 is connected to the T2 using a TSV, back side metallization layer (to enable TSV placement anywhere in the T1 die), and micro-bump on the top layer of the T2, that is then connected, through a series of metal layers of the T2, to the active component of the top tier (T2).

The benefits of using 3D-ICs are numerous and have been already pointed out in the literature very often over the past few years [8]. These advantages will be summarized in the next subsections.

1.3.1 3D-SIC advantages

Interconnection length

The 3D integration allows to design circuits with components closer to each other. Wire of a few millimetres long can be replaced by TSV of a few tens of microns, as shown in Fig. 1.3. These shorter interconnections will introduce shorter delays, hence allowing higher working frequencies [8, 13].

Silicon efficiency and accessibility

Adding a vertical dimension allows to increase the integration density. It is therefore possible to have more logic gates than a 2D-IC for the same footprint, hence a more efficient use of the silicon as shown in Fig. 1.4. For instance, compared to the

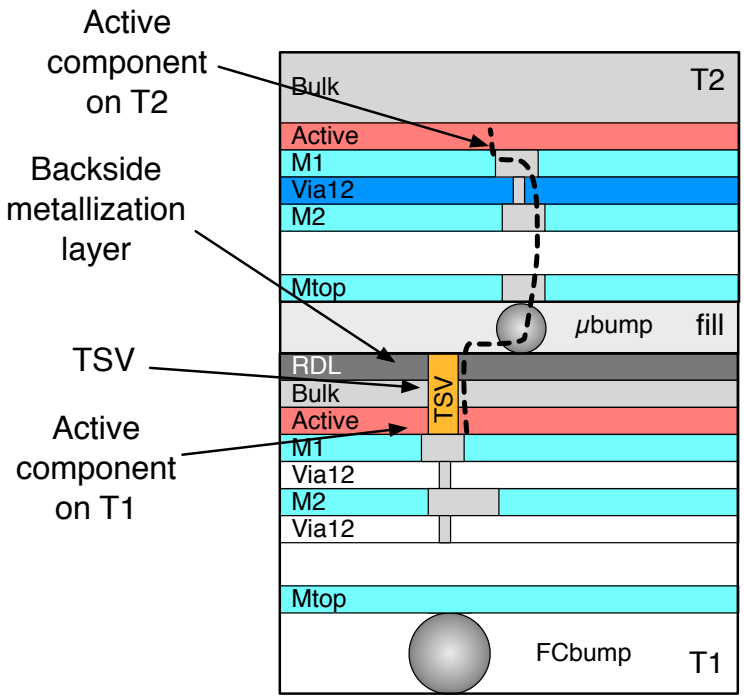

Figure 1.2: Illustration of the wiring properties of a 3D-SIC

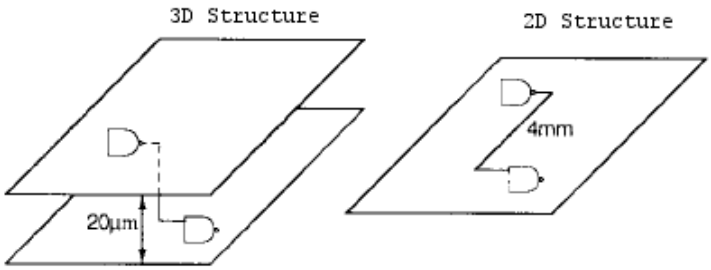

Figure 1.3: Shorter interconnections [8]

footprint of a 2D-IC, the 3D-SICs can double the integration for a 50% use of a 2D footprint [8].

In addition, the 3D integration allows a better accessibility for the components, as shown in Fig. 1.5. Indeed, for a 2D structure, 8 accessible neighbours can be considered for a central element (Fig. 1.5 (a)), whereas for a 3D structure, the number of accessible neighbours can reach 116 (Fig. 1.5 (b)) [8].

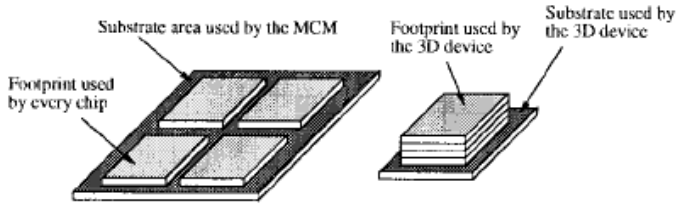

Figure 1.4: Silicon efficiency [8]

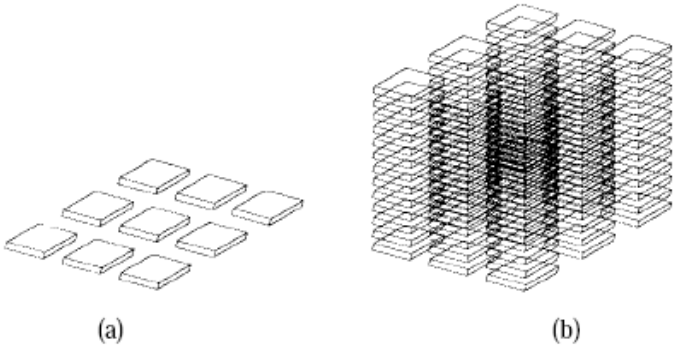

Figure 1.5: Components accessibility [8]

Bandwidth

The use of TSVs on 3D-SIC can significantly increase the bandwidth of a circuit. Indeed, as shown in Fig. 1.6, the interconnections are not only limited to peripheral connections but can also make use of the circuit's surface. This increase of the bandwidth allows higher working frequencies so that it is easier to satisfy data-heavy applications.

Consumption and noise

Shorter interconnections generally translates into lower capacitance and inductance parasites. This means a decrease of the numbers of repeaters, hence a better consumption, less noise and less jitter.

Heterogeneous circuits

The 3D technologies allow truly heterogeneous designs. For instance, it is possible to integrate, in addition to traditional digital circuits of different technologies, analogical circuits such as sensors or antennas, as well as power supply, which give

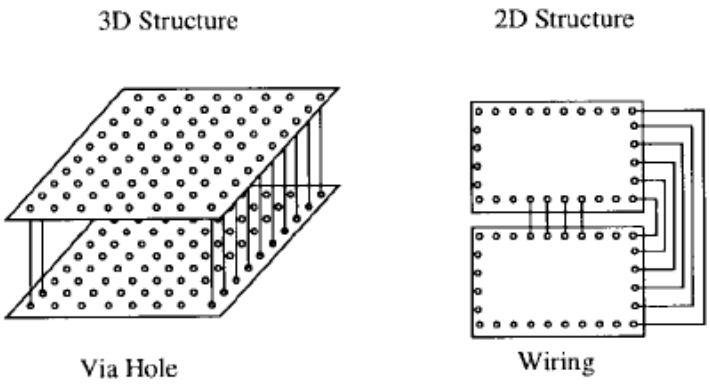

Figure 1.6: Bandwidth improvement [8]

3D-SIC a high degree diversity [14]. The Fig. 1.7 shows a schematic view of a 3D-SIC developed by IMEC for biomedical purposes that contains antennas, DSPs, EEG/ECG sensors, a power supply and solar cells [15].

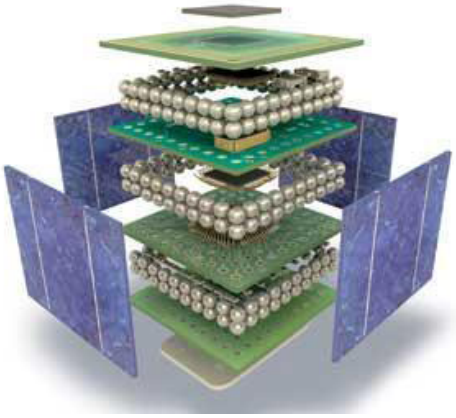

Figure 1.7: Schematic illustration of an heterogeneous 3D-SIC (developed by IMEC) [15]

1.3.2 Manufacturing technologies

Several 3D manufacturing technologies have been proposed and have been used to implement complete systems. Among the existing possibilities, four major categories of methods that illustrate 3D integration can be cited [8, 16].

Chip stacking

This methods consists in stacking components that have been designed and tested separately to produce a system-in-package (SiP). The vertically-stacked chips are interconnected with traditional wirings. The principal advantage of this method is an improvement in terms of size. The wirings are shorter however the components integration density is not increased compared to a 2D system.

Transistor stacking

The transistor stacking consists in creating several transistors level on one substrate. This should be the better way to manufacture 3D circuits although the success rate are currently limited due to thermal issues. The required temperatures to create a layer of high-performance transistors would provoke the destruction of the copper and aluminium already laid down on the previous layer.

Die-on-wafer stacking

In this method, known good dies (KGD), which are functional tested chips, are connected to a host wafer containing other KGDs. These KGDs can be interconnected with organic glues, oxide or metal bonding. The wafer and the bonded KGDs are then shaped to create the interconnections. Different substrates can be combined if the required temperature is low enough to minimize non-homogeneous expansion effects.

The die-on-wafer stacking can use interconnections on the edges of the chips or through-die. Depending on the interconnection type, this method can produce a better integration level than the chip stacking, with a better cost per connection ratio and a higher interconnection density, while holding the advantages of the KGDs.

The quality of the stacking depends on the pick-and-place equipment which is used to position the dies on the wafer. The placement accuracy will determine the possible interconnection density. Also, current equipments are supposed to handle fully buffered chips, not naked circuits so it does not provide protection to static discharge.

Wafer-level stacking

This methods consists in bonding entire wafers into a stack. The vertical through-wafer connections are made directly trough each substrate to the next wafer and it transistors layer. Similarly to the previous method, the interconnection density rely on the precision of the alignment, which is however currently better than the die-on-wafer stacking. This greater accuracy implies a better cost per connection ratio and a higher interconnection density compared to the die-on-wafer stacking.

The use of mixed substrates is also possible, only limited by the process temperatures. All the processing is done at the wafer level so wafer handling equipments are used. Since these provide protection to static discharge so there is no need to include buffering between the layers. The methods to bind two wafers are the same that are available for the die-on-wafer method.

One drawback to wafer-level stacking is its efficiency, since the chips on a wafer are not all KGDs.

1.3.3 3D-SIC design challenges

As explained, 3D-SICs offer numerous design perspectives thanks to their advantages. However there are drawbacks that need to be taken into account and that will be discussed in the following subsections.

Thermal dissipation

The power density has increased exponentially over the past decades for the 2D-ICs and it appears that this trend will continue in the near future. As for 3D-SICs, due to their higher component density, they will also be subject to higher power density so thermal management should be considered carefully [8].

Cost

With the appearance of a new technology, the involvement of a high cost should often be expected. In the case of 3D technology, the cost is currently high due to the lack of infrastructure and the reluctance of manufacturers who do not want to risk to change to new technologies [8].

Design complexity and design software

A large number of systems have been implemented using the 2D technologies which means that current tools can cope with 2D design complexity even if they show more and more their limits [4, 17]. As for 3D-SICs, the increased complexity can be tackled by developing adapted software [8]. However, to the best of our knowledge, few 3D dedicated software currently exist and they are mainly developed for and owned by particular manufacturers and are based on 2D design tools which does not allow to tackle the complexity of 3D designs integrally.

In the following section, we will have an overview about these software tools and generally about the design flow used to design integrated circuits;

1.4 Current design flows and their limitations

Design flows are the combination of electronic design automation (EDA) tools used to produce an integrated circuit. These flows can generally be summarized in 4 main steps [18], as shown in Fig. 1.8.

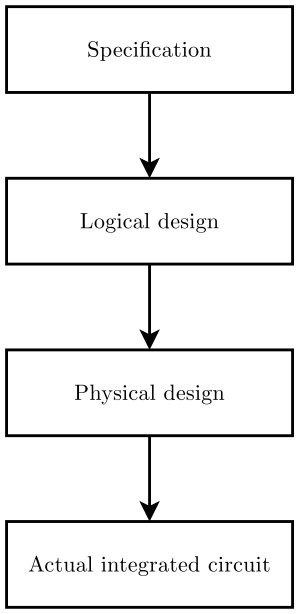

Figure 1.8: General classical design flow

As explained previously, designing ICs implies numerous choices. At the moment, with this growing complexity, the current design flows can already show their limits. For instance, most of the time, the designers will be likely to freeze a certain amount of choices on basis of their experience, and then begin the optimization process with the remaining parameters. This will therefore limit the exploration of the design space and good solutions may be ignored. In addition, the fixed choices can be questionable since they are based on the designer's experience though they could also be based on more objective facts.

The current design flows, which already showed their limits with conventional 2D-ICs, may thus need improvements to be able to deal with the increased complexity of emerging 3D-SICs [4, 17].

1.5 Design space exploration tools

In order to cope with the increasing complexity of integrated circuits and the limitations of the current design flows, numerous tools have been proposed, in particular works about design space exploration (DSE) that have been developed to quickly suggest possible solutions to a designer and speed up the design processes. In this section, we will describe different DSE tools that have been proposed in the literature.

1.5.1 2D-IC tools

MILAN

The MILAN (Model based Integrated simuLAtioN) framework [19] aims to simplify the optimization and the exploration of design spaces for SoC platforms. This tool works on the component level and allows the users to choose a compromise between the simulation speed and the results accuracy. The exploration and optimization process is done in two phases: first it searches for possible combinations between the architecture, the application and the mapping and second it estimates the performances (power, latency) depending the precision asked by the users.

SoC Architecture Explorer

SoC Architecture Explorer [20] is a multi-objective optimization and exploration tool that aims the design of SoC architectures by evaluating the compromises between the footprint and the execution time. The exploration process focuses on the application and the system architecture where the tool analyses the data flow and estimates the data transfers to determine a number of possible architectures.

modeFRONTIER (ESTECO)

modeFRONTIER [21] is a proprietary development environment developed by ESTECO. It is a multi-objective optimization tool that aims parallel SoC architectures. modeFRONTIER allows to deal with up to one million different design configurations thanks to statistical analysis tools and data mining techniques.

MULTICUBE

The MULTICUBE project (MULTI^3) [22, 23] is a European project started in 2008 and dedicated to the multi-objective exploration of MPSoC architectures for multimedia embedded systems. The aims is to developed a framework that allows a quick and automated exploration of the design space to improve the performances of a MPSoC with metrics such as power, latency, computing performance, bandwidth, QoS, etc. This project is based on several heuristics and optimization algorithms

that reduce the exploration time and allow a quick selection of the best solutions of a Pareto-optimal frontier. In addition, MULTICUBE also aims to define an application-oriented framework based on the results of the multi-objective exploration to optimize the resources allocation and the tasks scheduling of the applications. The exploration is done at the system level, using the SystemC language. The project includes proprietary and open-source tools whose development targets the industry. Among the developed prototyping tools, Multicube explorer and Multicube-SCoPE can be cited.

Multicube Explorer Multicube Explorer [24] is a design space exploration framework for supporting platform-based design. This tools allows a fast optimization of a system with objective functions such as power, delays, surface, etc. by means of a system simulator. Multicube explorer proposes several multi-objective optimization methods that aim to propose the best compromises.

Multicube-SCoPE Multicube-SCoPE [25] is an evolution of the SCoPE tool [26] oriented to design space exploration. It is a fast system performance and power simulator providing metrics associated with a system in order to drive the DSE process.

1.5.2 3D-SIC tools

DSE for 3D-stacked DRAMs by Weis *et al.*

DSE for 3D integrated circuits by Xie *et al.*

Automated design flow for 3D microarchitecture evaluation by Cong *et al.*

PathFinding Flow

1.6 Conclusion

To the best of our knowledge, all these tools use a uni-criterion approach or deal with a limited set of criteria while performing only trade-off analyses from a Pareto front. The goal of this research is to show that a more multicriteria-oriented optimization could be more suitable to take into account the many aspects of a design and provide more information.

In the next chapter, we will describe a short overview of the tools coming from the operations research that will allow to take into account multiple criteria simultaneously.

Bibliography

[1] R. X. Cringely. (2013) Breaking moore's law @ONLINE. [Online]. Available: http://betanews.com/2013/10/15/breaking-moores-law/

[2] S. Borkar, "Design perspectives on 22nm cmos and beyond," in *Design Automation Conference, 2009. DAC '09. 46th ACM/IEEE*, July 2009, pp. 93–94.

[3] V. Zhirnov, I. Cavin, R.K., J. Hutchby, and G. Bourianoff, "Limits to binary logic switch scaling - a gedanken model," *Proceedings of the IEEE*, vol. 91, no. 11, pp. 1934 – 1939, nov 2003.

[4] D. Milojevic, R. Varadarajan, D. Seynhaeve, and P. Marchal, *PathFinding and TechTuning in Three Dimensional System Integration*, A. Papanikolaou, D. Soudris, and R. Radojcic, Eds. Springer, Nov 2011. [Online]. Available: http://www.springer.com/architecture+%26+design/architecture/book/978-1-4419-0961-9

[5] S. Tans, A. Verschueren, and C. Dekker, "Room-temperature transistor based on a single carbon nanotube," *Nature*, vol. 393, no. 6680, pp. 49–52, 1998.

[6] Y. Cui, Z. Zhong, D. Wang, W. U. Wang, and C. M. Lieber, "High performance silicon nanowire field effect transistors," *Nano Letters*, vol. 3, no. 2, pp. 149–152, 2003. [Online]. Available: http://pubs.acs.org/doi/abs/10.1021/nl0258751

[7] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner, "Kondo effect in a single-electron transistor," *Nature*, vol. 391, no. 6663, pp. 156–159, Jan. 1998.

[8] S. Al-Sarawi, D. Abbott, and P. Franzon, "A review of 3-d packaging technology," *Components, Packaging, and Manufacturing Technology, Part B: Advanced Packaging, IEEE Transactions on*, vol. 21, no. 1, pp. 2 –14, feb 1998.

[9] E. Sicard and S. Dhia, "An illustration of 90nm CMOS layout on pc," in *Devices, Circuits and Systems, 2004. Proceedings of the Fifth IEEE International Caracas Conference on*, vol. 1, 3-5 2004, pp. 315 – 318.

[10] F. Microelectronics, "2008.1 product guide, assp, memory, asic," 2008.

[11] K. Takahashi and M. Sekiguchi, "Through silicon via and 3-d wafer/chip stacking technology," in *VLSI Circuits, 2006. Digest of Technical Papers. 2006 Symposium on*, 0-0 2006, pp. 89 –92.

[12] J.-S. Kim, C. S. Oh, H. Lee, D. Lee, H.-R. Hwang, S. Hwang, B. Na, J. Moon, J.-G. Kim, H. Park, J.-W. Ryu, K. Park, S.-K. Kang, S.-Y. Kim, H. Kim, J.-M. Bang, H. Cho, M. Jang, C. Han, J.-B. Lee, K. Kyung, J.-S. Choi, and Y.-H. Jun, "A 1.2v 12.8gb/s 2gb mobile wide-i/o dram with 4 #x00d7;128 i/os using tsv-based stacking," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International*, feb. 2011, pp. 496 –498.

[13] S. Das, A. Fan, K.-N. Chen, *et al.*, "Technology, performance, and computer-aided design of three-dimensional integrated circuits," pp. 108–115, 2004.

[14] E. Beyne and B. Swinnen, "3D System Integration Technologies," *Integrated Circuit Design and Technology, 2007. ICICDT '07. IEEE International Conference on*, pp. 1–3, May 2007.

[15] P. Pieters and E. Beyne, "3d wafer level packaging approach towards cost effective low loss high density 3d stacking," in *Electronic Packaging Technology, 2006. ICEPT '06. 7th International Conference on*, 26-29 2006, pp. 1 –4.

[16] 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.

[17] A. V. Biest, D. Milojevic, and F. Robert, "Key enablers for next generation system-level design in microelectronics," in *Proc. 10th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI), Orlando (Florida), 16-19 July 2006*, 2006.

[18] F. Robert, "Reconfigurable architectures," University Lecture, Université libre de Bruxelles, 2013.

[19] S. Mohanty and V. K. Prasanna, "Rapid system-level performance evaluation and optimization for application mapping onto soc architectures," in *in Proc. of the IEEE International ASIC/SOC Conference*, 2002.

[20] K. Ueda, K. Sakanushi, Y. Takeuchi, and M. Imai, "Architecture-level performance estimation method based on system-level profiling," *Computers and Digital Techniques, IEEE Proceedings* -, vol. 152, no. 1, pp. 12–19, 2005.

[21] ESTECO. (2001) Esteco - leader in engineering design optimization software, process integration and multidisciplinary optimization with modefrontier @ONLINE. [Online]. Available: http://www.esteco.com/index.jsp

[22] MULTICUBE. (2008) Multicube @ONLINE. [Online]. Available: http://www.multicube.eu/

[23] C. Silvano, G. Palermo, V. Zaccaria, W. Fornaciari, R. Zafalon, S. Bocchio, M. Martinez, M. Wouters, G. Vanmeerbeeck, P. Avasare, L. Onesti, C. Kavka, U. Bondi, G. Mariani, E. Villar, H. Posadas, C. Y. Q., F. Dongrui, and Z. Hao, "Multicube: Multi-objective design space exploration of multiprocessor architectures for embedded multimedia applications," in *Proceedings of the DATE'09 workshop on Designing for Embedded Parallel Computing Platforms: Architectures, Design Tools, and Applications*, Nice, France, April 2009.

[24] M. Explorer. (2009) Multicube explorer @ONLINE. [Online]. Available: http://home.dei.polimi.it/zaccaria/multicube_explorer_v1/Home.html

[25] Multicube-SCoPE. (2009) Multicube-scope @ONLINE. [Online]. Available: http://www.teisa.unican.es/gim/en/scope/multicube.html

[26] SCoPE. (2004) Scope @ONLINE. [Online]. Available: http://www.teisa.unican.es/gim/en/scope/scope_web/scope_home.php