A 3D reconfigurable platform for 4G telecom applications

Walid Lafi, Didier Lattard, Ahmed Jerraya
CEA, LETI, MINATEC
F38054 Grenoble, FRANCE
{firstname.lastname}@cea.fr

***Abstract*—To address the problem of prohibitive costs of advanced technologies, one solution consists in reusing masks to address a wide range of systems. This could be achieved by a modular circuit that can be stacked to build 3D systems with processing performance adapted to several applications. This paper focuses on 4G wireless telecom applications. We propose a basic circuit that meets the SISO (Single Input Single Output) transmission mode. By stacking multiple instances of this same circuit, it will be possible to address several MIMO (Multiple Input Multiple Output) modes. The proposed circuit is composed of several processing units interconnected by a 3D NoC and controlled by a host processor. Compared to a 2D reference platform, the proposed circuit keeps at least the same performance and power consumption in the context of 4G telecom applications, while reducing total mask cost.**

I. INTRODUCTION

Designing a custom integrated circuit is an option that offers significant benefits. The well-known rationale for incorporating a custom IC in a product is to reduce its manufacturing cost. In high volume, manufacturing a custom IC instead of a standard circuit is much less expensive. Moreover, a custom chip usually needs lower power than standard circuit, for a couple of reasons. Since a standard circuit is intended to address several applications that do not require the same energy consumption, applications with small power requirement may consume more than their real needs. Besides, many advanced hard-wired energy-saving techniques can be used in custom ICs and are not available when using standard circuits. For some applications such as mobile phones, small size and weight are crucial. A standard circuit is often much larger and heavier than custom chip. All these reasons make custom IC an attractive option for high-volume products.

Nevertheless, custom chips have higher non-recurring engineering costs than standard circuit. Indeed, a general VLSI application without regular system architecture requires multiple sets of masks. This can be extremely expensive since mask prices for cutting-edge processes have been increasing steadily. As shown in figure 1, the cost of only one mask set and its corresponding probe have already exceeded the 1 million euro. For this reason, reusing mask is becoming critical for advanced technology nodes.

978-3-9810801-7-9/DATE11/©2011 EDAA

In this work, our proposal is to design a modular circuit that could be stacked using 3D integration technologies in order to build 3D systems with processing performance adapted to several application requirements. Therefore, it would be possible to have a custom circuit for each application while using always the same mask set.

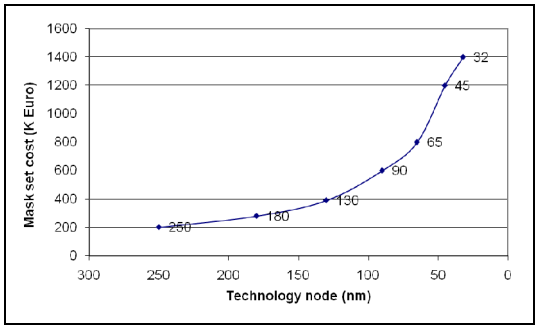

Figure 1. Increase in the cost of of mask set according to technological nodes

This paper is organized as follows. The proposed stackable circuit is presented in section II. Section III compares a 2-layer 3D system (built using our basic circuit) against the 2D reference platform in terms of performance and power consumption. Finally, section IV concludes the paper.

II. A STACKABLE & RECONFIGURABLE CIRCUIT FOR 4G TELECOM APPLICATIONS

In the case of mobile phones, the LTE standard (latest standard in the mobile network technology and often marketed as 4G) provides several transmission modes depending on the desired throughput and latency. As explained previously, it is preferable to have a custom circuit for each application. To address this issue, we present a reconfigurable NoC-based circuit for LTE applications. When used alone, the proposed circuit (henceforth called basic circuit) can meet the requirements of the SISO transmission mode. By stacking multiple instances of this basic circuit and performing some software reconfigurations, it will be possible to boost system performance and address several MIMO modes (figure 2). Let C be the mask set cost of an LTE telecom circuit in a given technology node. In order to have a custom circuit for each of the 4 LTE applications (SISO, MIMO2x2, MIMO4x2, MIMO4x4) with the classical 2D approach, a new mask set has to be designed for each application. Therefore, total mask

cost to address all the LTE applications is 4C. Using our 3D approach, only the mask set of the SISO transmission mode is needed to be designed. All other applications can be built using this same mask set. Total mask cost for all the LTE applications is only C. Therefore, by using the modular and stackable circuit, total mask cost could be reduced by 75%.

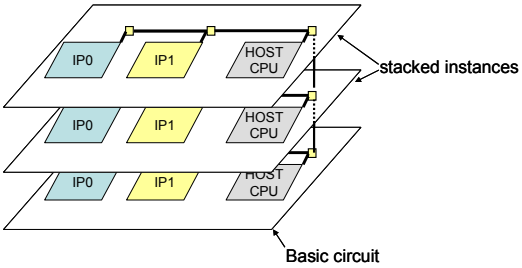

Figure 2. 3D reconfigurable circuit obtained by stacking multiple instances of a same basic circuit

In the rest of the paper, we prove through a rigorous case study, that in addition to this important gain in mask cost, this approach achieves at least the same performance and power consumption compared to a 2D reference circuit in the context of 3GPP LTE applications.

A. *Processing units*

To be able to satisfy the needs of several telecom applications, the basic circuit has to support data manipulation and data processing at the same time. To do so, three reconfigurable units are designed.

The first one is the Smart Memory Engine (SME) dedicated to memory management. It is a programmable unit designed to handle intensive data manipulation involving synchronization, buffering, duplication and reordering more efficiently than the data processing cores.

The second unit used in the proposed basic circuit is a digital signal processor called MEPHISTO. It is a high-performance reconfigurable core for complex matrices computation, useful for channel estimation, advanced MIMO coding/decoding… MEPHISTO is designed as a 32-bit data path Very-Long-Instruction-Word (VLIW) structure composed of a MAC (Multiplier/Accumulator) unit dedicated to complex arithmetic operations (16-bit I + 16-bit Q complex numbers), a compare/select unit for branch operations, and a cordic/divider unit for special computations.

The third component called TRX_OFDM is a configurable core for OFDM processing. It has a powerful complex FFT and IFFT datapath based on a butterfly implementation with an OFDM-symbol pipelined computation.

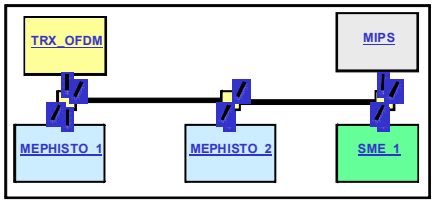

Figure 3. The basic circuit

Figure 3 depicts the resulting basic circuit. It is composed of 1 TRX_OFDM, 1 SME and 2 MEPHISTOs (to meet real time constraints) interconnected via 3 routers. Each unit is plugged on the NoC via a network interface (NI). The NI deals with packetization, depacketization and flow control using credits, handled by input/output communication controllers.

B. *MIPS-based semi-distributed control*

The previously described units require an efficient control mechanism to deal with scheduling and configuration. In this work, we use a semi-distributed control for the whole basic circuit. This allows alleviating the load of the host processor.

In addition to the local configuration and communication controller performed by the NI, a global control is performed by a 32-bit MIPS processor, by means of direct addressing and interrupts mechanisms. It is in charge of dynamic reconfigurations, real time scheduling and synchronizations. As depicted in figure 4, the MIPS processor has several extensions useful to interact and communicate with other basic circuit's components. These extensions include:

- an output extension managing the generation of data and configuration packets from the MIPS;
- an input extension allowing to read (to dump) data and configuration values from any of the circuit's units at the request of the MIPS;
- an interrupt controller in charge of handling interrupts generated in the NoC such as end-of-task notifications;
- and finally a local 16 KB RAM (32-bit word) used to store both instructions (the embedded control software), and data (configurations). An arbiter is used to allow the NoC to write the embedded control code in the MIPS's memory;

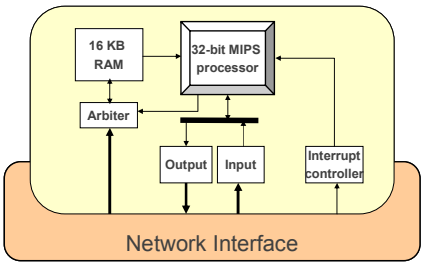

Figure 4. The MIPS-based global control unit

Unlike the local configuration and communication control performed by the NI, the global control is software-based to allow more flexibility. When multiple basic circuits are stacked (to build a 3D system), global control is distributed between all the MIPS processors of the resulting 3D stack. Each one is in charge of controlling the 4 components of its layer and communicates with other host processors to exchange information about scheduling. Compared to a centralized host processor approach, this approach allows distributing and then alleviating the global control load. Moreover, such a distributed approach improves scalability by avoiding control bottleneck problem when the number of processing cores increases.

C. *3D asynchronous mesh NoC*

All the units of the basic circuit are interconnected via a NoC. This NoC is also used to connect all the components of a

3D system resulting from stacking 2 or more basic circuits.

One of the most important design issues related to 3D circuit is global clock signal distribution to perform multi-plane synchronization [1, 2]. In order to avoid this problem, we propose to design the basic circuit as Globally Asynchronous Locally Synchronous (GALS) system. Processing units are synchronous (each one has its own clock frequency), while NoC routers are implemented in Quasi-Delay Insensitive asynchronous logic [3]. The NI performs synchronization between the synchronous and asynchronous domains. This asynchronous router achieves up to 18.7Gbit/s throughput. Its implementation details are not the scope of this paper.

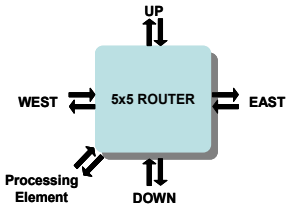

Figure 5. 5x5 asynchrounous router for intra-layer and inter-layer communications

In this work, we use only 5x5 routers (5 input ports x 5 output ports) to deal with all intra-layer and inter-layer communications as depicted in figure 5. The down ports of the routers located at the bottom layer are used to communicate with the external world.

D. Design results

The asynchronous router and all the previously described units were synthesised with 65nm low-power CMOS technology. All units designs include the NI and test mechanisms like scan chains and memory BIST. Table II depicts area results with the corresponding frequency of each synchronous core.

TSVs' area overhead is a key challenge limiting the viability of 3D circuits. The asynchronous router needs 184 links at each port to communicate with its neighbors. Considering high-density TSV with 4µm diameter and 10µm pitch (figure 6), total TSVs' area would be 12800µm^2, which corresponds to 6.5% of the router size.

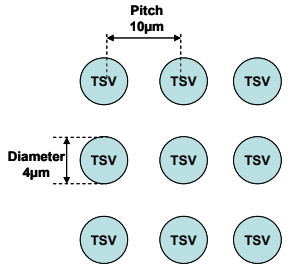

Figure 6. TSVs caracteristics

The basic circuit has a total area of 3.6mm^2. The TSVs' area overhead corresponds to 1.1% of the whole basic circuit size. We consider this overhead as negligible in this paper. The MIPS host processor induces a 4.8% silicon impact (<5%). The NoC infrastructure represents 17% of the whole circuit area.

TABLE I. SYNTHESIS RESULTS – 65NM TECHNOLOGY

Bloc	Frequency (MHz)	Area (mm^2)
MIPS	300	0.175
MEPHISTO	400	0.455
SME	400	1.274
TRX_OFDM	400	0.58
184 TSVs/router	-	0.0128
Router	-	0.20
Basic circuit	-	3.58
2D MAGALI	-	7.18

In this work, we take as 2D reference architecture a platform called MAGALI. This platform is devoted to wireless telecom applications. A silicon prototype is fabricated using the STMicroelectronics CMOS 65nm LP technology [4]. MAGALI platform focuses mainly on 3GPP LTE standard and aims multi antennas schemes (MIMO). MAGALI architecture consists of the same units and the same NoC used in our basic circuit. The MAGALI platform control (scheduling and configuration) is semi-distributed. A global control is centralized and performed by an ARM11 host processor. In order to make comparison with our 3D platform, we use only a sub-set of the MAGALI platform. This sub-set will be referred to as MAGALI in the rest of the paper.

The 3D system resulting from stacking 2 instances of the basic circuit is functionally equivalent to the MAGALI 2D platform. Thus, the 3D approach achieves up to 50% enhancement in form factor compared to the planar version.

III. PERFORMANCE AND POWER EVALUATION OF THE 3D PLATFORM

A. LTE benchmark application

To assess the performance and the power consumption of the proposed platform, we choose to deal with the downlink part of the LTE standard, and more specifically with the receiver side. Using the terminology defined in [5], data are transmitted in 10ms frames equally divided in 10 sub-frames also called TTIs (Time Transmission Intervals), i.e. the TTI aligns on sub-frame and equals 1 ms. The system is designed to transmit on 4 antennas and to receive on 2 antennas (4x2 MIMO), which requires a high performance processing.

In this work, we consider a 2-layer 3D system resulting from stacking 2 instances of the basic circuit. The targeted application is mapped to this 3D platform. To have a relevant comparison, an equivalent mapping is performed on the 2D reference architecture MAGALI.

B. Performance results

Simulation environment includes 2 SystemC data generators to emulate the incoming data-flow from the 2 antennas. To speed up simulation, the asynchronous NoC is modelled in TLM SystemC with post-layout parameters. All units of the 2 platforms are modelled at RTL level to provide cycle-accurate results.

Table II summarizes the performance results corresponding to the processing time of a complete TTI including scheduling and reconfiguration phases. Host processors of the 2 platforms run at 300MHz clock frequency, while processing units run at 400MHz.

As depicted is table II, the execution time is almost the same in the two platforms. This is expected since we are using the same processing units on the 2 platforms. From 2D MAGALI to the 3D 2-layer system, the scheduling and reconfiguration management are transferred from centralized host CPU to 2 distributed MIPS processors. By efficiently using the smaller MIPS processor, there is no time overhead due to the communication between the 2 control processors of the 2 layers.

The results based on RTL simulation, confirm that the 3D platform with a distributed control is as efficient as the 2D MAGALI platform with a centralized controller.

TABLE II. TIME TO PROCESS A TTI

Platform	2D MAGALI	3D platform
Execution time	456µs	454µs
performance speed up	-	0.4 %

C. *Power consumption results*

To provide a full comparison, we have evaluated the power consumption of the 2 platforms. Each platform was placed and routed in 65nm low-power CMOS technology. A complete TTI processing was simulated with the placed and routed netlist.

Table III presents the average power consumption of the different processing units of the 2 platforms at gate level. The contributions of the FFT, data processing and data manipulation are the same for the 2 platforms since we are using the same units.

TABLE III. POWER CONSUMPTION OF THE PROCESSING UNITS TO PROCESS A TTI

	2D MAGALI / 3D platform (mW)
FFT	172
Processing	207
Data manipulation	258
Total	637

The ARM11 host processor of the MAGALI platform deals not only with control, but also with the MAC (Medium Access Control) layer. Dealing with the MAC layer is not the scope of this paper. Therefore, it would not be relevant to make comparison between ARM11 and the MIPS host processors (which are devoted only to control) in terms of the power consumption. Table IV depicts the average power consumption of the MIPS control processor corresponding to different activity rates. With 0% activity rate, the MIPS processor consumes 14.9 mW corresponding to the clock tree power consumption. By efficiently programming the MIPS processor, it is active only 0.4% over the time of TTI period to perform scheduling and reconfiguration, and its power consumption reaches 19mW. Therefore, it is possible to achieve an important power consumption gain (more than 50%) by inserting a simple clock gating mechanism in the MIPS processor. Figure 7 depicts a detailed view of the power consumption of the MIPS processor profile during a TTI processing.

These results based on post place & route simulation, prove that the distributed control approach introduce a low power consumption overhead to perform the control of the whole 3D platform.

TABLE IV. POWER CONSUPTION OF THE CONTROL PROCESSORS

	Activity rates	Power consumption
MIPS without clock gating	0%	14.9mW
MIPS without clock gating	0.4%	19mW
MIPS with clock gating	0.4%	9mW

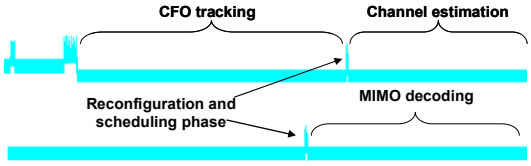

Figure 7. Power consuption profile of the up-layer MIPS procesor

IV. CONCLUSION

In this paper, we present a reconfigurable circuit for 4G telecom applications able to meet the requirements of the SISO mode of transmission. Thanks to its reconfigurable units, stacking multiple instances of this basic circuit allows boosting system performance and addressing several MIMO modes. By designing the modular and stackable circuit, it is possible to use always the same mask set and then to reduce total mask cost. A rigorous comparison (based on synthesis, RTL simulations and post place & route simulations) is performed between a 2-layer 3D system and the 2D MAGALI reference platform. The 3D approach shows at least the same performance and power consumption of the processing units.

Future work will focus mainly on enhancing power management (by means of dynamic task mapping for example), in order to deal with thermal constraints that could limit the viability of our approach.

REFERENCES

[1] V.F Pavlidis. I. Savidis, E.G. Friedman, "Clock Distribution Networks for 3-D Integrated Circuits", *IEEE Custom Integrated Circuits Conference, 2008 (CICC 2008)*, 2008 , pp. 651- 654.

[2] M. Mondal et al., "Thermally Robust Clocking Schemes for 3D Integrated Circuits", *Proceedings of the IEEE International Conference on Design, Automation and Test in Europe*, pp. 1206-1211, April 2007.

[3] Y. Thonnart, P. Vivet and F. Clermidy, "A Fully-Asynchronous Low-Power Framework for GALS NoC Integration", *IEEE International Conference on Design, Automation and Test in Europe*, Mars 2010.

[4] F. Clermidy, C. Bernard, R. Lemaire, J. Martin,I. Miro-Panades, Y. Thonnart1, P. Vivet, N. Wehn, "A 477mW NoC-Based Digital Baseband for MIMO 4G SDR", *In Proceeding of IEEE International Solid-State Circuits Conference (ISSCC) 2010*, Volume 53, pages 278-279, February, 2010, San Francisco, USA.

[5] 3GPP TSG-RAN, 3GPP TS36.211, *Physical Channels and Modulation (Release 8)*, v8.1.0 (2007-11).