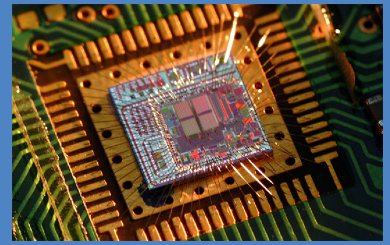# IcySoC

INEXACT SUB-NEAR-THRESHOLD SYSTEMS FOR ULTRA-LOW POWER DEVICES

Prof. Andreas P. Burg,
EPFL

Prof. Luca Benini,
ETHZ

Prof. Christian Enz,
EPFL

Philippe Rochaix,
EM Microelectronic

David Ruffieux,
CSEM

What it's about…

Developing an ultra-low-power platform based on an integrated circuit operated at very low supply voltage ("near- or sub-threshold") and using inexact computation blocks that provide approximate results tolerated by many applications like video or audio.

Context and project goals

The notion of exact computation, where outputs of the computational element (circuit) have precise deterministic values, as well as the fact that electronic chips are powered at nominal voltages for increased performances, have been pervasive in the computing domain for many decades owing to the overwhelming success of the integrated circuit design using reliable transistors, particularly in Complementary Metal-Oxide-Semiconductor (CMOS) technology. However, semiconductor industry is facing serious challenges today as diminishing transistor sizes driven by Moore's law are leading to increasing process variations and additional perturbations due to temperature and voltage fluctuations which threaten the circuit functionality. Owing to such widely anticipated hurdles to continued technology scaling - the promise of Moore's law - and a growing desire for reducing energy consumption, techniques and technologies such as inexact/approximate circuits and sub- or near-threshold circuits (supply voltage below or near the transistor threshold voltages) have gained prominence. The first radical approach realizes parsimonious or "adequately engineered" designs that trade accuracy at the hardware level for significant gains in energy consumption, area, and speed. The second approach offers the minimal power or energy consumption at the cost of increased delay and power variations. A large class of energy constrained systems, particularly in the domain of embedded portable multimedia and in domains of budding interest such as recognition, search and data mining, lend themselves readily for such a design philosophy. In fact, all of which can tolerate inaccuracies to varying extents or can synthesize accurate (or sufficient) information even from inaccurate computations.

Until now, these research works have been limited to application-specific instances of building blocks that were mostly ad-hoc targeting some specific examples and did not consider well-understood complete platforms based on these inexact and extreme low voltage components in sub- or near-threshold operation. In addition, research was conducted without a synergy between inexact computing and extreme low voltage circuits. It is therefore mandatory to consider at the same time the design of various inexact, approximate, sub- or near-threshold components and the platform consisting of these components. The platform design will be largely impacted by the usage of these components, in terms of parallelism, performances and robustness. One has to revisit the system design in terms of usage of hardware accelerators, heterogeneous or homogeneous processor cores and of communication or network-on-chip that has to be implemented for data transmission.

It has been demonstrated that inexact arithmetic blocks could provide a reduction up to 15X in delay, power and area product. Sub- or near-threshold circuits could provide a reduction of 6X in dynamic power when reducing the supply voltage from 1.0 V. to 0.4V. The platform design, while using very energy-efficient hardware accelerators, will contribute to the significant power reduction expected from the combination of the aforementioned techniques.

We will address practical issues by using the proposed techniques to fabricate prototype chips implementing large-scale error resilient systems and through physical measurements to validate and demonstrate evidence of the utility of these techniques both quantitatively (through well-defined application-specific quality metrics) and qualitatively, yielding perceptually discernible outputs (such as audio, image or video data).

How it differentiates from similar projects in the field

The project combines three techniques that are nicely complementary: first, a multiprocessor platform (parallelization is mandatory to reduce power while keeping the same throughput), second extreme low supply voltages down to 0.3 or 0.4 Volt (as dynamic power is proportional to Vdd2, significant power reduction) and third, inexact arithmetic that also reduce significantly the power consumption.

Quick summary of the project status and key results

The team is very close to tapeout a complete platform with 4 cores, memories, exact and inexact hardware accelerators in ALP180 (180 nm from EM Microelectronics). Sub-/Near-threshold Standard cell libraries in 180 nm and 65 nm have been developed, as well as memories. Inexact arithmetic adders and multipliers have been designed using two different approaches: pruning at gate level and speculative adders. Dynamic RAMs (DRAMs) and low-voltage standard-cell based memories (SCMs) based on latches in different technologies (180 nm, 65 nm, and 28 nm) have been developed.

Development of techniques and circuits for timing error detection and for highly dynamic clock-frequency adjustment of microprocessors that adjust their clock period on a cycle-by-cycle basis based on current instructions. Integration of many chips in 65 nm comprising four cores and various exact and inexact hardware accelerators like FPUs.

Success stories

The platform chip, a quite complex chip in ALP180, comprising 4 cores, memories, exact and inexact hardware accelerators, to tape out in June 2015, integrated in summer and available in September for measurements.

In addition, the standard cell libraries and memories in ALP180 and 65 nm have been released. Several inexact blocks have been proposed based on two different approaches: gate-level pruning and speculative adders. Dynamic RAM memories and low power memories have been proposed and integrated in 65 nm.

Microprocessors augmented with timing error detection and for highly dynamic clock-frequency adjustment have been proposed.

Main publications

G. Karakonstantis, A. Sankaranarayanan, M. M. Sabry, D. Atienza, A. Burg, A Quality-Scalable and Energy-Efficient Approach for Spectral Analysis of Heart Rate Variability, DATE 2014, Dresden, Germany, March 24-28, 2014.

A. Lingamneni, C. Enz, K. Palem, C. Piguet, Highly Energy-efficient and Quality-tunable Inexact FFT Accelerators, CICC 2014, September 15 – 17, 2014, San Jose, California.

J. Constantin, L. Wang, G. Karakonstantis, A. Chattopadhyay, A. Burg, Exploiting Dynamic Timing Margins in Microprocessors for Frequency-Over-Scaling with Instruction-Based Clock Adjustment, DATE 2015, March 9-13, 2015, Grenoble, France.

A.Teman, G. Karakonstantis, R. Giterman, P. Meinerzhagen, A. Burg, Energy versus Data Integrity Trade-Offs in Embedded High-Density Logic Compatible Dynamic Memories, DATE 2015, March 9-13, 2015, Grenoble, France.

V. Camus, J. Schlachter, C. Enz, Energy-Efficient Inexact Speculative Adder with High Performance and Accuracy Control, ISCAS 2015, May 24-27, 2015, Lisbon, Portugal.

J. Schlachter, V. Camus, C. Enz, K. V. Palem, Automatic Generation of Inexact Digital Circuits by Gate-level Pruning, ISCAS 2015, May 24-27, 2015, Lisbon, Portugal.

V. Camus, J. Schlachter, C. Enz, Energy-Efficient Digital Design through Inexact and Approximate Arithmetic Circuits, NEWCAS, June 7-10, 2015, Grenoble.

S. Ganapathy, A. Teman, R. Giterman, A. Burg, G. Karakonstantis, Approximate Computing with Unreliable Dynamic Memories, NEWCAS, June 7-10, 2015, Grenoble.

"Oh, that's near enough!"

Letting microchips make a few mistakes here and there could make them much faster and more energy-efficient