**11.** **Publishable summary: "Reliable and Variability tolerant System-on-a-chip Design in More-Moore Technologies"**

Project Facts:
- FP7 Project :   European Community funded
- Coordination :   IMEC
- Website :   www.fp7-reality.eu
- Duration :   30 Months
- Effort :   382 person-months
- Industry :   ARM (UK), ST Microelectronics (Italy)
- Start date :   1st January 2008
- University :   Glasgow (UK), Bologna (Italy), Leuven (Belgium)
- Research Centre :   IMEC (Belgium)

Scope:
- Scaling beyond the 32 nm technology
- Tackle the increased variability and changing performance of devices from device unto system level.
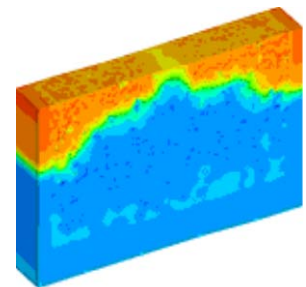
Random discrete dopants in a 35 nm MOSFET from the present 90 nm technology node.

Challenges:
- Increased static variability and static fault rates of devices and interconnects.
- Increased time-dependent dynamic variability and dynamic fault rates.
- Build reliable systems out of unreliable technology while maintaining design productivity.
- Deploy design techniques that allow technology scalable energy efficient SoC systems while guaranteeing real-time performance constraints.
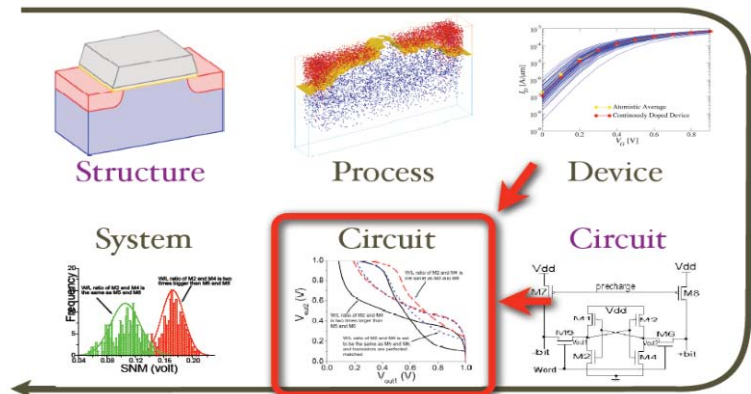
Proposed solution:
- System analysis of performance, power, yield and reliability of manufactured instances across a wide spectrum of operating conditions.
- Generally applicable solution techniques to mitigate the impact of reliability issues of integrated circuits, at component, circuit, and architecture and system design.
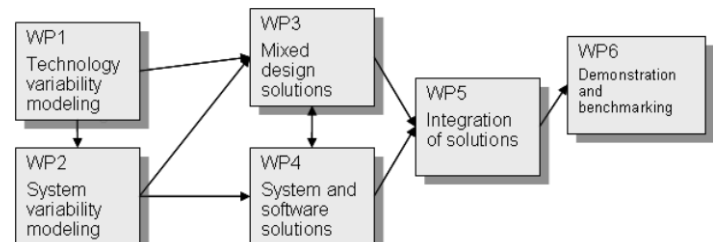
WP1: Device variability and Reliability Models (WP leader: UoG)

Task T1.1.1 (Validation and calibration against 45 nm technology generation devices) was completed and reported in 2008. This year's efforts have been focused on tasks T1.1.2 (started in 2009) and tasks T1.2.1, T1.2.2, T1.3.1, T1.3.2 and T1.3.3 already started in 2008.

T1.1.2   Simulation of variability in 32 nm technology generation devices.
An NDA was signed on 9/4/2009 to allow transfer of 32nm TCAD data from ST to UoG with the 4-way agreement being signed on 15/5/2009. The TCAD data was transferred to UoG on 28/04/2009. UoG has introduced the provided doping profiles into their simulator and calibrated the simulations against the provided I-V curves. UoG has performed simulations to investigate the variability in threshold voltage due to random discrete dopants and line edge roughness. A new source of variability (workfunction variation due to metal gate granularity) has also been included in the simulator for the first time. Full I_D-V_G characteristics for both n- and p-channel devices, at high and low drain voltage, have been simulated. Ensembles of 200 devices that differ due to random discrete dopants, line edge roughness and metal gate granularity have been used.

T1.2.1   Statistical simulation of the impact of fixed/trapped charges.
Simulation of the impact of fixed/trapped charges in 45nm devices was completed and reported in 2008. Further analysis of the simulation results was performed in early 2009 and the results have been submitted for publication in IEEE Transactions on Electron Devices. The equivalent work for 32nm devices will be performed at the start of 2010.

T1.2.2   Modelling of time dependent variation:
Simulation models for NBTI, HCD and SDB: this was completed and reported in 2008.
Variability aware method to incorporate reliability effects (NBTI, HCD and/or SBD) in transistor level netlists: This task goal is a modelling technique to include reliability effects in a variability aware transistor level model of any given circuit with the goal of simulating the combined impact of static variability and reliability. The inputs are the extracted variability injectors from 3.3 and the simulation models for reliability defined the previous year in 2.2

T1.3.1   Statistical compact model parameter sets.
45nm statistical compact models were delivered by UoG and reported in 2008. Due to differences in release number between the 32nm compact model supplied by STM and the TCAD simulation data to which the UoG simulator was calibrated, a different approach has been adopted for the 32nm devices. UoG have provided the I_D-V_G characteristics from T1.1.2 to IMEC who have extracted variability injectors that can be used in the provided STM compact models. This allows other partners to use compact models with variability without violating the NDA between STM and UoG on the supply of device information.

T1.3.2   Extraction of key statistical parameters.
More physically based compact models, such as PSP, will allow the extraction of key parameters which could have geometry dependent statistical distributions: UoG has developed a statistical extraction strategy for PSP.

T1.3.3   Extraction of variability injectors
Method and tool flow to extract Vt and beta variations from I-V curves (device DC transfer characteristics): Application of extraction flow to ST 45nm data was completed and reported in 2008. Application of extraction flow to ST 32nm data to extract Vt and beta variations from I-V curves provided by UoG (based on the low-level technological information received from ST) has been done in 2009 and variability injectors have been provided to project partners.

WP2: System and circuit characterization and sensitivity analysis (WP leader: IMEC)

Although it might be considered a straightforward task, statistical library characterization is a critical showstopper to the adoption of statistical analysis and optimization techniques in industrial digital flows. ARM has kept on developing with his EDA partner the characterization tool capable of generating a CCSA-VA liberty file with a high accuracy. A good correlation between analysed variability and Monte Carlo simulation was finally achieved. A 80 cell library has been characterized and will be evaluated on a system level design. In parallel, ST developed and validated a practical, yet accurate technique to reduce the mismatch characterization time during the reporting period. It is based on the simultaneous identification of the transistors impacting the cell performances, on the reduction of the characterization grid size, and on the identification of a restricted set of transistor parameters that largely impact on the device performances, since it would not be feasible to consider all the parameters present in typical transistor compact models.

While industry support exists for library characterization of standard cells, entirely new ways have to be gone for non-standard cells. ARM developed a push button tool for the designers to easily apply the read current characterization, based on spreadsheet tables and SPICE. In parallel, implications of Silicon-on-Insulator (SOI) technology on variability are investigated. KUL has prepared a memory for use in imec's MemoryVAM characterization tool.

On digital block level, ST introduced its new hybrid approach. A traditional DSTA based on corners is carried out using the tools currently available in the digital sign-off flow. After DSTA, ST considers the critical paths violating the timing constraints and also those paths that are "potentially" critical, i.e., whose timing slack are very small and subject to the detrimental impact of process variability. During this analysis, OCVs are taken into account by means of derating factors (which can be overly pessimistic as it was demonstrated during the previous WP2 activities). After selecting the critical/violating paths, on one side ST reduces the OCV margins (based on process variability data) and also carries out a path-based SSTA (HSTA) based on the statistical libraries described earlier. Following this approach it is possible to remove timing violations without taking unnecessary margins.

Imec has continued to implement and verify the technique to translate IP-level variability to System-level variability. The electronic Information Format (IF) that imec develops to exist parallel to the classical top-down and bottom-up design flow is continuously extended in order to cope with advanced information like block-level-sensitivity data.

WP3: Mixed mode countermeasures (WP leader: KUL)

The work of KUL between M13 and M24 has focused on task T3.2.

Task T3.2 has been divided into two sections: a first section extends the work done in T3.1 to include variability awareness to the reliability simulation of electronic circuits. In a second section, this knowledge is used to design a variability/reliability-resilient memory IP block.

A. Variability and degradation phenomena demand the use of novel design techniques, e.g. Knobs & Monitors, to guarantee optimal performance over a specified lifespan. A nominal reliability simulation methodology is used in combination with a variability-aware framework to maintain simulation accuracy while taking process variations into account. The methodology has been demonstrated effectively, but efficiency can still be improved in the coming months.
B. The design of a memory IP-block has been carried out as a case study to develop and validate variability-aware and low-power design methodologies for a state-of-the-art memory. To be able to cope with variability, statistical behavioural models are required. In this way the number of design loop iterations between the architectural and the circuit level can be reduced. The techniques that have been studied in this work are: (fully) divided word and bit lines, charge recycling, asymmetric sense amplifier redundancy, low swing techniques, etc. This work has temporarily been halted due to the lack of 32nm device models, which has caused serious delays and personeffort underspending. Part of this task has also been done collaboratively by IMEC, who set up the VAM environment and ran the statistical characterization flow on the KUL SRAM memory.

Since the work in this task has not yet been completed due to the delay and personeffort underspending as discussed above, KUL will continue and finalize this work by extension of Task T3.2 into June 2010. The completion of this work is also essential to support the activities in WP6. KUL will report this extended work in an updated version D3.2b of the Deliverable D3.2.

The work of UNIBO between M13 and M24 has focused on T3.4, and more specifically analyzing variability effects on on-chip communication and developing techniques for compensating for them and supporting run-time adaptation and variability management. The analysis has identified on-chip communication links (and more specifically Network-on-chip links) as critical components to be analyzed and protected by adequate countermeasures, based on Adaptive body bias (ABB) and adaptive supply voltage (ASV), which have been identified in T3.3 as effective methods for post-silicon tuning to reduce variability on generic combinational circuits or microprocessor circuit sub-blocks. The results have been reported in D3.3 (delivered on time at M12).

WP4: System level countermeasures (WP leader: UNIBO)

The work performed by UNIBO until M18 was focused on Task 4.1, 4.2, 4.3 and 4.4 (in collaboration with ST). The first two tasks (results are reported in D4.1 – delivered on time at M9) allowed to build the software infrastructure for the development of variability countermeasures developed in Task4.3 (results are reported in D4.2 – delivered on time at M18).

Main activities related to Task4.3 have been:

1. Definition of the problem of task allocation under variability effects to minimize energy under time constraints. We devised an Optimal solution with Integer Linear Programming. It is very slow and suitable for off-line validation. We then developed a sub-optimal solution based on Linear Programming and BinPacking, which is suitable for implementing task allocation at startup time. Finally, we developed an approximate solution based on a Look-up Table, which is much faster and allows online implementation.
2. Development of a theoretical solution and its implementation on the simulation platform that has been enhanced with power and variability models as part of WP5
3. Refinement of the task model (independently, barrier synchronized tasks representative of embarrassingly parallel computation like as done in video processing algorithm)
4. Comparison against state-of-the-art policies: Rank Power/Rank Freq
5. Analysis of compensation capabilities using VAM-generated variability affected platforms.

The activity about Task 4.4 started in collaboration with ST. In particular, UNIBO is currently working on the optimization of the policies on the target multicore platform, provided by ST and on the target validation benchmarks defined in WP6, for which UNIBO is also performing the porting to the simulation platform as part of WP6.

The activity performed by UNIBO during the period M19-M24 in WP4 was focused on the completion of Task 4.4, concerning the optimization of the system level countermeasures for the specific hardware platform defined in WP5 and the benchmarks to be used for evaluation in WP6. More specifically, the activities have been:

1. The development of an optimized online version of the variability-aware task allocation policy exploiting a new formulation of Linear Programming problem that can be solved during application execution. This optimization is needed for the implementation of the task allocation policy on multimedia applications on a frame-by-frame basis.
2. The analysis of the impact of variability on the performance and energy consumption of multimedia applications, namely on the MPEG2 video decoder benchmark from ST (WP6). The application is parallelized so that some functions (i.e. IDCT and Decode Slice) are performed by various cores at the same time. This analysis provided information about the impact of speed and power consumption variations among the different cores on the energy and deadline miss rate.
3. The implementation of the optimized variability-aware allocation policy on the target platform running the MPEG2 video decoding benchmark.

More details and results related to this activities will be reported in D4.3 due on M27.

WP5: Design flow, integration, proof of concept (WP leader: ARM)

This Work Package coordinates and carries out the evaluation of the different project components. It is the place where the components of the project come together and will involve developing methodologies and test platforms to enable the evaluation to take place. It is responsible for assessing the severity of the un-compensated situation, and quantifying the improvements that result from the application of techniques identified in other Work Packages. It is also responsible for identifying the techniques that prove to be unsuccessful (and why), as well as to point to other techniques that might be worth investigation (beyond this project).

In a first phase we defined the system platforms and the IP blocks in which we will evaluate the developed techniques. Some initial developments on the xSTream and the ARM system platforms have already started in the first period.

The second period has seen a ramping up of the integration activity. The flow integration started on the first period on several cells has been extended to a full library. The work has been focused on the VAM flow from IMEC. A 32nm library composed of 80 cells is used as bases of the validation work. A full variability analysis has been performed on this library. Thanks to the deployment of a Research Island inside ARM a strong cooperation on the variability analysis of the VAM flow was undertaken. IMEC could hereby get access to the ARM926 design and the 32nm library.

In the scope of the variability characterization a ARM memory critical path has been characterized for different memory cut sizes. One size was chosen to perform a benchmarking of memory design coming out from the WP3 developments. Memory VAM was also applied to the memory developed by KUL.

The ALU of the ARM926 was chosen as the test vehicle for the Active Body Biasing scheme evaluation. UniBO started the adaptation of its methodology using a 45nm library provided by ST. The 32nm library, including back biasing information, became available at the end of the period allowing testing this technique on the 32nm node as well.

On the xSTream platform an important effort has been spent in the development of the simulation infrastructure of the system level design. An adaptation of an in house flexible model has been done in order to be able to simulate the xSTream processor and the xPE IP blocks including variability information. The variability impact seen on the system will be modelled by the "back of the envelop" approach proposed by UniBo. This approach is based on equation providing the relationship between corner variations and the system performances. A calibration of the variability information will be done using the VAM information on the ARM926. In order to account for the variability information at the system level during the simulation, specific adaptations of the software infrastructure had to be made.

Finally a monitoring system has been integrated to the xSTream platform. The system is based on a library component: Monitoring Micro IP. The information coming out from the IP can be accessed either by a Process Monitoring block which controls the knobs on the system or directly read out during the testing of the system.

This second period has seen a lot of activity in this work package. Most of the integration work has been done and is described in the delivery D5.3 which paves the path to the assessments done in WP6.

WP6: Validation and assessment of results (WP leader: ST)

This work package is the place were most of the results, methodologies, flows, IPs developed and/or leveraged in the other work packages come together for validating the project's goal of "build reliable systems with unreliable components".

The WP tasks are designed to validate and assess the project results through the benchmarking of individual IP blocks along with the identification and porting of a set of industrially relevant applications. The general focus is the leverage of IPs and applications along with their integration and mapping onto the integrated application platform developed with contributions from WP3,4,5. The output of the work package is the evaluation both in terms of NRE, final product cost, and in general advantages over the uncompensated case, of the benefit stemming from the methodologies, techniques and components developed by the project.

During the last six months of the second year of execution of WP6 activities, the work has progressed mainly for tasks T6.4, T6.5.1 and T6.5.2, allocation of resources to task has been aligned to expected levels from

the DoW. Tasks T6.2 and T6.3 have been completed during the first six months of this year and reported in the previous Periodic Activity Report.

In the context of WP6 validation tasks, as reported in the last Periodic Activity Reports, the partners of REALITY have agreed to validate the technologies and methodologies developed with two separated flows for the multimedia (STM) and general purpose (ARM) platform components.

The ARM926 test case has been applied towards the validation and benchmarking of both variability aware circuits and design and analysis flows; the process was driven by the availability of 32nm silicon technology data complemented by advanced variability analysis techniques at the atomistic level and up to devices, standard cell libraries, circuits and memories and whole netlists.

The major output of the period is the release of the report associated to D5.3; in this report are detailed many of the validation and benchmarking activities and results carried out by the partners on the different IPs and analysis flows deployed in the project.

For the multimedia scenario, a suitably defined simulation platform based on the REALITY platform definition tasks of WP5 has been made available. Such simulation platform provides a set of parallel media acceleration engines from STM (called xPEs) and an STM proprietary host processor, ST231, supporting Linux and a suitable run-time environment for the accelerators.

Objectives for the period 2, Project M13 until M24

The reporting Period 2 covers the project time-schedule M13 until M24, i.e. starting from 1st January 2009 until 31st December 2009. This report is a mid year progress report for such period.

Description of the performance / research indicators (all targets for Y2 have been met!)

WP	After year 1	After year 2	At end of project
WP1 Device variability	Physical modeling and understanding of the variability at 45/32 nm technology nodes (TN). First statistical compact models.	Compact models that accurately capture the variability and the reliability issues at 32 nm.	Models fine tuned. Feedback from device measurements incorporated.
	Preliminary version of a RDR std. cell library [32nm]. Flow definition and framework set up for variability characterization. Correlated variability energy timing flow definition and set up.	Variability characterization of a [32nm] RDR std cell library. Exploitation of the variability aware modelling flow on the driver application vehicle, including a solution for IP blocks and memories.	Methodology fine tuned. Feedback from benchmarking acknowledged.
WP3 Mixed design	Description of the variability and reliability analysis methods at circuit level	Demonstration of the developed method on SRAM and analog circuits	Validation of the developed method
WP4 Algorithm	Software techniques for flexible data and workload allocation for migration (the base flexible RTSM support)	Control algorithms for system level reliability and variability management (exploiting of the base RTSM support)	Porting, optimization and tuning for the target evaluation platform of: (a) the flexible RTSM, (b) the control algorithms
WP5 Integration	Definition of characterization blocks, macrocells, and system level architecture	Validation and application of methods to macrocells and integration into system	Final system integration, validation feeding into WP6 benchmarking
WP6 Benchmarking	Identification of relevant industrial applications and associated requirements and evaluation metrics	Definition of the validation plan Benchmarking of block level IPs	Benchmarking of system level platform Evaluation of results and impact according to validation plan criteria