Robust Modular Bulk Built-In Current Sensors for Detection of Transient Faults
Frank Sill Torres, Rodrigo Possamai Bastos

To cite this version:

HAL Id: lirmm-00715122
https://hal-lirmm.ccsd.cnrs.fr/lirmm-00715122
Submitted on 6 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Robust Modular Bulk Built-In Current Sensors for Detection of Transient Faults

Frank Sill Torres
Dept. of Electronic Engineering
Federal University of Minas Gerais
Belo Horizonte, Brazil
franksill@ufmg.br

Rodrigo Possamai Bastos
LIRMM
Université Montpellier II / CNRS UMR 5506
Montpellier, France
bastos@lirmm.fr

Abstract—Soft error resilience is an increasingly important requirement of integrated circuits realized in CMOS nanometer technologies. Among the several approaches, Bulk Built-in Current Sensors (BBICS) offer a promising solution as they are able to detect particle strikes immediately after its occurrence. Based on this idea we demonstrate a novel modular BBICS (mBBICS) that tackles the main problems of these integrated sensors – area, leakage, and robustness. Simulations based on a predictive nanometer technology indicate competitive response times for high performance applications at the cost of 25 % area overhead and very low power penalty. Thereby, all simulated particle strikes that lead to transient faults could be detected. Additionally reliability analysis proved the robustness of the proposed mBBICS against wide variations of temperature and process parameters.

Keywords: Built-in current sensors, concurrent detection, fault tolerance, reliability, security, soft errors, transient faults

I. INTRODUCTION

CMOS is furthermore the most widespread technology for integrated designs as no feasible alternative is in sight to date and in the near future. Driving forces of this leadership are the high miniaturization capability and the reliability of CMOS. Against the background of nanotechnology though, reliability concerns are arising with an alarming pace. The shrinking of technology sizes result in circuits that are susceptible to several errors sources like oxide breakdown [1], parameter variations [2] or radiation [3]. In case of the latter, energetic particle can inject electrical charge into sensitive regions of the semiconductor devices and, thus, lead to soft errors. Until the early 2000’s, researches on this kind of errors focused mostly on memories and application oriented on avionics and aerospace environment. In current nanometer technologies though, soft error resilience is also required for ground level applications and the combinational parts of the circuits. In recent years, several concurrent error detection and/or correction techniques have been proposed to mitigate the effects of soft errors. This includes the use of debug resources from the scan path [4], the combination of error detecting codes with carry-save arithmetic [5], and selective redundancy [6]. In contrast to the before mentioned approaches which focus on gate and system level, Bulk Built-In Current Sensors (BBICS) offer a promising solution on transistor level to detect particle strikes immediately after its occurrence [7]. The advantages of this technique are zero delay penalty and fast error detection.

On the downside, existing BBICS are prone to temperature and parameter variations and/or require a considerable amount of area [8]. Further, several known BBICS implementations increase the leakage power dissipation of the monitored circuits.

The aim of this work is to present a new modular BBICS (mBBICS) which operates reliably under wide temperature and process variations, has reasonable area requirements, and has low impact on power dissipation.

The rest of the paper is organized as follows: section 2 gives basic information about the content of this work. Section 3 describes the concept of the proposed sensor while section 4 is related to simulation results and discusses the findings. Finally, section 5 draws the conclusion.

II. FUNDAMENTALS

This section presents preliminary information regarding soft errors, Bulk Built-In Current Sensors (BBICS) and leakage currents in nanometer technologies.

A. Transient Faults in Integrated Circuits

Transient voltage variations during the lifetime of combinational or sequential circuits are defined as transient faults. The first harmful effects of transient faults are soft errors by inverting stored results of system operations (i.e. bit-flips).

Due to the transistor shrinking and the growing communication of confidential data, soft errors can happen today even at ground level by means of perturbation events arisen from environmental or intentional sources. Examples of environment events are alpha particles released by radioactive impurities and more importantly neutrons from cosmic rays [7]. On the other hand, intentional perturbation events are usually produced by optical sources such as flashlights or laser beams [8] that can maliciously induce transient effects on secure circuits like smartcards to retrieve their secret information.

Soft errors and transient faults are also known respectively as single event upsets (SEU) and single event transients (SET) in fault-tolerance-related fields or even as consequences of attacks based on fault injection in security applications.

A transient fault in CMOS circuits is modeled by injecting a double exponential current pulse Ifault at the sensitive node. Thereby, the shape of the pulse can be approximated by following equation [9]:

This work was sponsored by PRPq/UFMG, FAPEMIG, CAPES, and CNPq.
with \( Q_f \) is the charge collected due to the particle strike, \( t_f \) means the decay time of the current pulse, and \( t_r \) labels the time constant for initially establishing the ion track.

### B. BICS Detecting Transient Faults

Built-in current sensors (BICS) were initially proposed as a mechanism for detecting large increases in the current \( I_{DDQ} \) consumed by a CMOS circuit during its quiescent state, i.e. when the circuit is not switching. The mechanism allows thus testing CMOS circuits against permanent faults [10]. Further, BICS were also adapted for detecting transient faults in memory cells (i.e. bit-flips) [11-13]. Recently, efforts were made for monitoring transient faults in combinational logic as well [14]. All these techniques connect BICS to the power lines (VDD and GND) of the monitored circuit in order to distinguish anomalous transient currents from normal currents. The today’s problem is that the amplitude of transient currents induced by radiation effects or fault attacks can have the same order of currents normally generated by switching activities in combinational logic circuits. Hence, schemes monitoring power lines are very limited for detecting just small range of transient faults.

On the other hand, BICS connected to the bulks of the monitored circuit’s transistors are able to detect a wide range of transient faults [7, 8]. As Fig. 1(a) and (b) illustrate, Bulk-BICS (BBICS) identify anomalous transient currents \( I_A \) flowing through the junction between a bulk and a reversely-biased drain of a perturbed transistor (MOSFETs “off” in Fig. 1). BBICS indeed take advantage of two facts: (1) in fault-free scenarios (i.e. \( I_A = 0 \)) the bulk-to-drain (or drain-to-bulk) current \( I_B \) is negligible even if the MOSFET is switching in function of new input stimuli; and (2) during transient-fault scenarios \( I_A \) is much higher than the leakage current flowing through the junction. The range of detectable transient faults is easily adjustable by calibrating the size of the transistors that constitute the BBICS. Hence, schemes based on BBICS can be designed to latch a flag of fault indication for abnormal currents within a defined range that represents a risk of resulting in soft errors.

Fig. 1 (c) summarizes the strategy for protecting system’s blocks against transient faults in pull-down network by using BBICS. Equivalent strategy must be taken for detecting transient faults in pull-up network as well. Note that in such a strategy the connection between the monitored circuit (e.g. system’s block 1) and the BBICS’s circuitry (e.g. NMOS-BBICS 1) is done via metal – from the body-ties of each monitored transistor (e.g. T11, T12, and T13) up to the input of the BBICS’s circuitry. Thereby, the peak of the anomalous transient current (i.e. the transient fault) is almost not attenuated, ensuring thus an efficient detection. In fact, this very small attenuation is a function of the local distance between the struck zone of the monitored transistor and its body-tie.

### C. Leakage

Ideally, CMOS cells draw no current or rather dissipate no power when idle. Unfortunately, this is not true for real cells realized in nanometer technologies. A major impact originates from sub-threshold leakage current \( I_{sub} \) which is the current between source and drain when the transistor is in theory cut off. A commonly used approximation of \( I_{sub} \) is [15]:

\[
I_{sub} = a_t \frac{W}{L} e^{\frac{q}{a_d} \left[ \frac{V_{gs}}{n V_{th}} - \frac{V_{th}}{n V_{th}} \right]} \tag{2}
\]

with \( a_t \) is a technology factor, \( W \) means the gate width, \( L \) labels the gate length, \( q \) corresponds to the charge of an electron, \( n \) means the sub-threshold swing coefficient, \( k_B \) is Boltzmann’s constant, \( T \) labels the operating temperature, \( V_{gs} \) is the gate-source voltage, and \( V_{th} \) labels the threshold voltage. Further, the threshold voltage can be modeled with [15]:

\[
V_{th} = V_{th0} + \gamma a_d V_{sh} - \eta a_d V_{ps} \tag{3}
\]
with $v_{th0}$ is the zero-bias threshold voltage, $\gamma$ is the body-bias coefficient, $a_2$ and $a_3$ label technology constants, $\eta$ corresponds to the is Drain Induced Barrier Lowering (DIBL) coefficient, $V_{sd}$ labels the source-bulk voltage, and $V_{ds}$ is the drain-source voltage.

III. MODULAR BBICS

In this section, the underlying assumptions as well as the proposed sensor structure is presented.

A. Origination of the approach

Existing BBICS are capable of detecting smallest bulk currents with low response times. Moreover, the approach is much more efficient for dealing with transient faults of long duration and multiple faults without impact on the system’s operating frequency. However, common problems are high area effort as well as strong susceptibility to parameter and temperature variations. Further, most available BBICS create area effort as well as strong susceptibility to parameter and operating frequency. However, common problems are high duration and multiple faults without impact on the system’s functionality between different BBICS as possible solution to reduce the required area overhead. Further, we decided to connect the bulk via a high ohmic transistor in on-state with VDD or GND, respectively, as proposed in [17]. Last but not least, we searched a positive feedback structure to increase stability and decrease the sensor response time.

Based on these observations, we investigated on solutions for reducing the sensor area, increasing the sensor robustness, and avoiding any voltage offset on the bulk at normal operation (i.e. in fault-free scenarios). Thereby, we identified sharing of functionality between different BBICS as possible solution to reduce the required area overhead. Further, we decided to connect the bulk via a high ohmic transistor in on-state with VDD or GND, respectively, as proposed in [17]. Last but not least, we searched a positive feedback structure to increase stability and decrease the sensor response time.

Next, the basic structure and the mode of operation of the proposed modular BBICS, in the following named as mBBICS, is presented. The considerations are related to the sensor for the NMOS bulk. The PMOS version, whose behavior is complementary, is omitted for the sake of simplicity.

B. Basic Structure

Figure 2 depicts the basic structure of the proposed mBBICS which can be divided in two kinds of components: head and tail. Thereby, several heads which are connected to separate bulks share one tail.

A head connects via transistor Nh1 the bulk of the Block Under Test (BUT) with GND. The transistor which has a small W/L ratio is on-state. The bulk is also connected with the gate input of transistor Nh2 whose drain connects to the common signal $head_{NMOS}$ of several heads.

A tail which is similar to an asynchronous latch consists of an inverter realized via the transistors Pt1 and Nt1 whose input is connected with $head_{NMOS}$. Its output $inv_{out}$ feeds the gates of the devices Pt2 and Nt2, as well as the input of the inverter INV3. The output of the latter is the error signal $error_{NMOS}$. Further, the output of the inverter formed by Pt2 and Nt2 connects to $head_{NMOS}$ which is also connected to VDD via transistor Pt3 whose gate is controlled by the signal reset.

C. Mode of Operation

In normal operation, i.e. no particle strike occurs, the bulk is at GND level (at $\odot$ in Figure 3). Further, the net $head_{NMOS}$ has VDD potential and the devices Nh2 in the head modules are in off-state. In case of a particle strike within the BUT, a current flows through transistor Nh1 in the related head module. This results in a voltage drop over Nh1 based on its channel resistance. Consequently, the bulk level increases (at $\oplus$ in Figure 3) and Nh2 starts to conduct leading to a reduction of the $head_{NMOS}$ voltage level (at $\ominus$ in Figure 3).

In the tail block, the decrease of the level of $head_{NMOS}$ activates the inverter formed by Pt1 and Nt1, whereby the state of signal $inv_{out}$ starts to change from GND to VDD (at $\odot$ in Figure 3). As consequence, transistor Pt2 changes from on-state to off-state, and Nt2 vice versa, which in turn decreases the level of $head_{NMOS}$ and, thus, forces the change of $inv_{out}$ to VDD level. This positive feedback stops when $inv_{out}$ reaches VDD level while $head_{NMOS}$ is at GND level. At the same time, the error signal $error_{NMOS}$ changes to VDD level (at $\ominus$ in Figure 3), which can be processed by higher instances as a flag indicating the occurrence of a transient fault. Consequently, the mBBICS has to be reset after such a process by changing Pt3 shortly into on-state (at $\odot$ in Figure 3) and, thus, setting the level of $head_{NMOS}$ back to VDD level. It has to be observed that when a flag of fault is latched and during the reset of the mBBICS no new transient fault can be detected.

Figure 2: Structure of the proposed mBBICS (version for NMOS bulk)
It should be observed that based on the described positive feedback and adequate sizing, which will be explained in the next subsection, it is not required that the head circuit reduces the signal head_NMOS below VDD/2. Thus, variations of the head circuit characteristics have lower impact on the functionality of the mBBICS.

D. Sizing

The devices Nh1 in the head modules must have a low W/L ratio (i.e. L is longer than W) to guarantee sufficient on-resistance. In contrast, the devices Nh2 must be sized with a great W/L ratio to assure that already minor changes of its gate-source voltage $V_{gs,Nh2}$ result in considerably drain-source current $I_{ds,Nh2}$ that changes the level of head_NMOS. As long as Nh2 is in its sub-threshold region, i.e. $V_{gs,Nh2}$ is smaller than the threshold voltage of Nh2, $I_{ds,Nh2}$ is equal to the sub-threshold leakage (see eq. 2). Hereby, the sensor profits from the high sub-threshold leakage of integrated circuits realized in nanometer technologies.

The devices Nt1 and Pt1 in the head module must be sized to support a GND-to-VDD change of the net inv_out. Hence, Pt1 should have a high W/L ratio and Nt1 a small one. In a first guess, the transistors Pt2 and Nt2 of the following inverter should be sized to support a VDD-to-GND transition of the signal head_NMOS. This is true for Pt2. However, during our analysis of the mBBICS behavior in nanometer technologies we could observe that the leakage through Nt2 is critical as it is discharging the node head_NMOS. Therefore, Nt2 should be minimum sized, or even with a greater gate length which considerably reduces the leakage (see also eq. 1). Device Pt3 as well as the inverter can be minimum sized.

IV. RESULTS AND DISCUSSION

This section presents the attained findings and discusses its implications.

A. Simulation Environment

All simulations are based on predictive 16 nm technology models using a VDD of 0.7 V [20][21]. Each Block Under Test (BUT) consists of six chains of ten inverters, whereas each inverter is sized with a driving strength of two. The bulk of all NMOS devices of a BUT is connected to one head circuit. In contrast, the bulk of the PMOS devices of each BUT is monitored by two head circuits whereas each head is connected to the bulk of the half of the PMOS transistors. This configuration assures equal capacitive load of the monitored PMOS and NMOS bulk. The transistor sizes of the modules of each mBBICS type are listed in Table I.

A particle strike was simulated by a current pulse based on eq. 1 at the output node of the 5th inverter of the first chain, in the following named sensitive node. All simulation were executed with a rise time $t_r$ of 1 ps.

B. Nominal Case

At first, we determined the susceptibility of the BUT to particle strikes. Therefore, we varied the collected charge $Q_f$ and the decay time $t_f$ (see also eq. 1). Then, we measured the voltage peaks at the sensitive node and the output of the inverter chain. We defined a transient fault TF when the voltage at the sensitive node as well as at the output of the inverter chain crossed VDD/2.

Next, we estimated the ability of the proposed mBBICS sensor to detect TF if the parameters of the applied technology have its nominal values. Based on initial robustness approximations we chose an mBBICS configuration with 6 heads and one tail circuit for the PMOS and NMOS version as reference implementation. The simulations were executed for the same $Q_f$ and $t_f$ values as in previous susceptibility analysis. The results, including the response time $t_{resp}$ are shown in Table II whereas $O\times$ – marks that no TF and no detection occurred, $O\checkmark$ – marks that no TF but a detection occurred, $\star \star$ marks that a TF but no detection occurred, and $\star \star \checkmark$ – marks that a TF and a detection occurred. It can be seen that all transient faults could be detected by the mBBICS. However, for the highest charge a false detection happened for the NMOS as well the PMOS mBBICS. This case is unfavorable though, its impact on the design performance is bearable as it results only a small amount of additional clock cycles. Further follows that the maximum response time is 452 ps for the PMOS mBBICS and 430 ps for the NMOS version.

In the following step we analyzed the response time of the NMOS and PMOS mBBICS in function of the number of heads. Thereby, the simulations were executed for the case $Q_f = 2 \, \mu C, t_f = 5$ ps which was determined as lower border for the occurrence of a transient fault. The results as well as the area overhead are depicted in Figure 4. The response time

---

**TABLE I. TRANSISTOR SIZES IN NM FOR THE APPLIED MBBICS WHEREAS THE 1ST NUMBER INDICATES THE WIDTH AND THE 2ND THE LENGTH (THE OMITTED INVERTER INV3 IS MINIMUM Sized)**

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>name</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nh1</td>
<td>16 / 160</td>
<td>Ph1</td>
<td>16 / 176</td>
</tr>
<tr>
<td>Nh2</td>
<td>160 / 18</td>
<td>Ph2</td>
<td>224 / 18</td>
</tr>
<tr>
<td>Pt1</td>
<td>64 / 16</td>
<td>Pt1</td>
<td>16 / 160</td>
</tr>
<tr>
<td>Nt1</td>
<td>16 / 64</td>
<td>Nt1</td>
<td>160 / 16</td>
</tr>
<tr>
<td>Pt2</td>
<td>16 / 160</td>
<td>Pt2</td>
<td>16 / 32</td>
</tr>
<tr>
<td>Nt2</td>
<td>16 / 32</td>
<td>Nt2</td>
<td>16 / 320</td>
</tr>
<tr>
<td>Pt3</td>
<td>16 / 16</td>
<td>Nt3</td>
<td>16 / 16</td>
</tr>
</tbody>
</table>

---

![Figure 3: Voltage curves of a mBBICS (NMOS version) detecting a transient fault](image-url)
TABLE II. DETECTION CAPABILITY OF PROPOSED MBBCS (6 HEADS, O - NO TF, NO DETECTION; • • - NO TF, DETECTION, • • • - TF, DETECTION) AND RESPONSE TIME $t_{resp}$

<table>
<thead>
<tr>
<th>$Q_f$</th>
<th>$t_f$</th>
<th>PMOS $t_{resp}$</th>
<th>NMOS $t_{resp}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 fC</td>
<td>5 ps</td>
<td>O •</td>
<td>•</td>
</tr>
<tr>
<td>2 fC</td>
<td>5 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
<tr>
<td>3 fC</td>
<td>5 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
<tr>
<td>4 fC</td>
<td>5 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
<tr>
<td>1 fC</td>
<td>10 ps</td>
<td>O •</td>
<td>• •</td>
</tr>
<tr>
<td>2 fC</td>
<td>10 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
<tr>
<td>3 fC</td>
<td>10 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
<tr>
<td>4 fC</td>
<td>10 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
<tr>
<td>1 fC</td>
<td>20 ps</td>
<td>O •</td>
<td>• •</td>
</tr>
<tr>
<td>2 fC</td>
<td>20 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
<tr>
<td>3 fC</td>
<td>20 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
<tr>
<td>4 fC</td>
<td>20 ps</td>
<td>• •</td>
<td>• •</td>
</tr>
</tbody>
</table>

The response time in that range varies for the PMOS MBBCS between 157 ps and 598 ps.

The area overhead starts with 49 % for MBBCS with only one head circuit and settles at around 21 % for the maximum supported number of heads. Thereby, for MBBCS with six heads the area overhead reduced already to 25 %.

In the last step, we determined the impact on delay and power dissipation in normal operation, i.e. no particle strikes.

Robustness Analysis

Robustness is a critical requirement for integrated sensor as variations of all kinds are continuously increasing. Therefore, the main impact on the sensor performance comes from temperature and process variations.

For the first robustness analysis we analyzed the response time as well as the functionality of the MBBCS in function of the temperature. All simulations were executed for the reference configuration with 6 heads, a collected charge of $Q_f = 2$ fC and a decay time of $t_f = 5$ ps. The results are depicted in Figure 5. We could observe that for temperatures below -15 °C (PMOS MBBCS) and 0 °C (NMOS MBBCS), respectively, the transient fault could not be detected. Further, the simulations showed that for temperatures above 45 °C (PMOS MBBCS) and 60 °C (NMOS MBBCS), respectively, false detection occur. Hence, both MBBCS versions work in a range of 60 °C.

In the second robustness analysis we executed MonteCarlo simulations utilizing a normal distribution for the varied technology parameters. The expected values $\mu$ and the standard deviations $\sigma$ for the modified BSIMv4 parameters are depicted in Table III [15]. For each configuration we executed 50 MonteCarlo simulations which is recommended amount [19].

At first, we determined the functionality and maximum response time in dependence of the collected charge for the reference MBBCS configuration of 6 heads. The results are depicted in Figure 6 and indicate a maximum response time of 1091 ps (PMOS MBBCS) and 778 ps (NMOS MBBCS) each for collected charge of 2 fC. Further, we could observe that all

varies for the NMOS MBBCS between 157 ps and 598 ps.

The response time in that range varies for the PMOS MBBCS between 157 ps and 598 ps.

The area overhead starts with 49 % for MBBCS with only one head circuit and settles at around 21 % for the maximum supported number of heads. Thereby, for MBBCS with six heads the area overhead reduced already to 25 %.

In the last step, we determined the impact on delay and power dissipation in normal operation, i.e. no particle strikes.

We estimated the maximum delay of the inverter chains with $t_{max}$ for the NMOS MBBCS and 22 for the PMOS version. A higher number leads to more false detections is the sub-threshold leakage of the Nh2 transistors in the head circuits.

The area overhead starts with 49 % for MBBCS with only one head circuit and settles at around 21 % for the maximum supported number of heads. Thereby, for MBBCS with six heads the area overhead reduced already to 25 %.

In the last step, we determined the impact on delay and power dissipation in normal operation, i.e. no particle strikes. We estimated the maximum delay of the inverter chains with $t_{delay}$ for the modified BSIMv4 parameters are depicted in Table III [15]. For each configuration we executed 50 MonteCarlo simulations which is recommended amount [19].

At first, we determined the functionality and maximum response time in dependence of the collected charge for the reference MBBCS configuration of 6 heads. The results are depicted in Figure 6 and indicate a maximum response time of 1091 ps (PMOS MBBCS) and 778 ps (NMOS MBBCS) each for collected charge of 2 fC. Further, we could observe that all
in all simulation all transient faults could be detected and that in no case a false detection occurred if the no particle strike happened. For a collected charge of $Q_f = 1 \text{ fC}$ the NMOS mBBICS showed no false detection while the PMOS mBBICS had 6.

Finally, we determined the functionality and maximum response time for greater amount of chains. Thereby, we could observed that up to 8 heads both mBICCS types never suffered under false detections if no particle strike occurred. The maximum response time increased to 1418 ps (PMOS) and 1384 ps (NMOS).

V. CONCLUSIONS

Bulk Built-In Current Sensors (BBICS) offer a promising solution to cope with soft error problems in current nanometer technologies. However, this kind of sensors suffers under its susceptibility to temperature and parameter variations, as well as area and power increase. The proposed modular BBICS applies functional block sharing and positive feedback to mitigate these problems. Simulations with a predictive 16 nm technology indicated that the proposed sensor could detect all injected transition faults with response times below 500 ps for the nominal case, and 1 ns under wide process and temperature variations. This could be achieved with a reasonable area offset of 25 % and very low increase in power dissipation.

VI. REFERENCES


![Figure 6: Maximum response time in function of collected charge $Q_f$ (6 heads, $t_f = 5$ ps, 50 MonteCarlo simulations)](image-url)