Thermal Stress Monitoring using Gradient Direction Sensors

Ahmed Lakhssassi1 Ahmed.lakhsasi@uqo.ca
Mohammed Bougataya1,2 bougmo01@uqo.ca
Charbel Boustany1 bouc29@uqo.ca
Daniel Massicotte2 Daniel.massicotte@uqtr.ca


Index Terms—FPGA, VLSI, Thermal Analysis, Sensors, GDS technique.

Abstract:

During the development of a VLSI (Very Large Scale Integration) circuits; their internal stress due to packaging combined with local self heating becomes serious and may result in large performance variation, circuit malfunction and even chip cracking. Surface peaks thermal detection is necessary in large VLSI circuits. This paper presents a VLSI circuit thermal stress monitoring approach using surface peak thermal detector algorithm and GDS (Gradient Direction Sensors) method. The design of surface peak thermal detector algorithm (SPTDA) with flexible modular-based architecture will be presented. Several approaches were implemented to achieve a better performance for the SPTDA algorithm operation. A parallel processing strategy is used to minimize computational delay. Furthermore, a hardware-efficient factoring approach for calculating tangent and division functions required by SPTDA algorithm is used to minimize silicon space in regards of their implementation. Description of the algorithm developed for the surface peaks thermal detection and the architecture implementation results are reported and compared with finite element method (FEM) temperature computations.

I. INTRODUCTION

Decreasing feature sizes and increasing power and package contact densities are making thermal issues extremely important in large VLSI design. Thermal monitoring is a crucial vehicle for predicting the change in the electrical characteristics or possible stress-induced failure of a large VLSI system. The rapid evolution in the industry of the integrated circuits during the last decade was so quick that currently it is possible to integrate complex systems on a single SoC (System on a Chip). Due to aggressive technology scaling, VLSI integration density as well as power density increases drastically. For example, the power density of high performance microprocessors has already reached 50W/cm² at 100 nm technology and it will reach 100W/cm² at 50nm technology [1]. This evolution towards higher integration levels is motivated by the needs of advanced high performance, lighter and more compact systems with less power consumption. Meanwhile, to mitigate the overall power consumption, many low power techniques such as dynamic power management [2], clock gating [3], voltage islands, dual V_DD/V_RL and power gating, are proposed recently. These techniques, though helpful to reduce the overall power consumption, may cause significant on-chip thermal gradients and local hot spots due to different clock/power gating activities and varying voltage scaling. It has been reported in [3] that temperature variations of 30°C can occur in a high performance microprocessor design. The magnitude of thermal gradients and associated thermo-mechanical stress is expected to increase further as VLSI designs move into nanometer processes and multi-GHz frequencies.

The principal effect of the absence of a good dynamic thermal management is the gradual and continuous degradation of the quality of performance as well as some other direct effects on the life cycle of the electronic systems [4]. This paper presents a VLSI thermal stress monitoring approach, using GDS, for development of SPTDA algorithm. The proposed algorithm using only two sensor cells will be formulated in a manner to facilitate the development of modular architectures using minimum silicon space in regards of their implementation in VLSI. The architecture selected will be modeled in high level languages, simulated in order to evaluate its performances, and then implemented on a FPGA (Field-Programmable Gate Array). A closed loop of simulation is used in order to evaluate the performances of the architectures proposed at each stage. The proposed architecture of the algorithm will be designed in a modular perspective after the separation of the different elementary functions of the algorithm. Hence the design of SPTDA with flexible modular-based architecture will be presented. The architecture is designed in high level languages such as Matlab™ - Simulink®, simulated, tested using VHDL and synthesized using Xilinx™ ISE [5] and Altera™ DSP [6] tools. The simulation and hardware implementation results obtained will be compared to a finite element method (FEM) temperature computation of the entire GDS method configuration cells.

II. GRADIENT DIRECTION SENSORS METHOD

The proposed algorithm is based on the GDS method for evaluating a single heat source on the chip surface. This method principle of work is explained in details in [7]. For two sensors, A and C placed in the distance a (Fig. 1), the difference between their output voltages is proportional to the changes of the temperature value along the distance a [7]. This is true only when the heat source is directly on the
line AC for any other cases the values of the angle $\alpha$ has to be taken into account for the proper calculation of $\Delta T$ (1).

$$\frac{\Delta T}{\Delta r} = \frac{T_c - T_d}{a \cdot \cos \alpha} \leftrightarrow \frac{V_c - V_d}{a \cdot \cos \alpha}$$

(1)

where $r$ is the distance from the heat source.

In figure 1, we have:

$$b = AD \cdot AD \leftrightarrow T_b - T_d \leftrightarrow V_b - V_d$$

(2)

$$b + c = AE \cdot AE \leftrightarrow T_c - T_d \leftrightarrow V_c - V_d$$

(3)

In order to obtain the information about angle $\alpha$ we should apply the third sensor. In the simplest case the GDS contains only three temperature sensors placed in distance $a$ (fig. 1).

Figure 2 shows the proposed distribution of the 6 sensors divided into 2 cells located whether on or outside the chip.

### III. DESIGN METHODOLOGY

#### III-A Presentation of algorithmic division

The general flowchart of the proposed SPTDA algorithm is shown in figure 3. It shows some parallelism in computation between the internal variables represented by $\tan \alpha_1$ and $\tan \alpha_2$. This parallelism will be used later in the architectural design in order to optimize the speed of the running implementation. Since successive evaluation of heat source temperature is needed for fast heat-source changes. The algorithm will create 2 triplets of values coming form the detectors. These 2 triplets will be used to calculate the tangents values and estimate $R_1$, $R_2$ and the temperature of the heat source. By dividing the algorithm into its basic arithmetic elementary operations (AEO) we will be able to formulate its data flow graph in order to model the algorithm.

![Flowchart of the SPTDA algorithm](image)

On the basis of eqs. (2) and (3) and the geometrical dependencies from (fig. 1), eq. (4) is obtained:

$$\tan \alpha = \frac{2}{\sqrt{3}} \left( \frac{V_b - V_d}{V_c - V_d} - \frac{1}{2} \right)$$

(4)

Using the cell we can obtain information on the temperature distribution and partly on the position of the heat source. In order to obtain the temperature value of a single punctual heat source we have to calculate the distance between the sensor and this source. Two sensor cells are required for this purpose (fig. 2). Two sensor cells A, B, C and D, E, F are placed in two corners of a monitored layout in the distance $H$. Hence, the temperature of the heat source can be obtained by equation (5).

$$T_s \leftrightarrow V_s = \frac{H}{a} (V_c - V_d) \cdot \frac{(\tan \alpha_1 + 1) \cdot (\sqrt{3} + \tan \alpha_2)}{\sqrt{3} \cdot (1 - \tan \alpha_1 \cdot \tan \alpha_2) - (\tan \alpha_1 + \tan \alpha_2)} + V_s$$

(5)

**Fig. 1** The 3 sensors cell $\alpha \in (0^\circ, 30^\circ)$ [7]

**Fig. 2** Problem description and the distribution of the sensors cells

On the basis of eqs. (2) and (3) and the geometrical dependencies from (fig. 1), eq. (4) is obtained:

**Fig. 3** Flowchart of the SPTDA algorithm

**Fig. 4** show that a quantification combination of a signed 11.4 and 12.4 bits had great differences with the main function; 13.4 was the turning point, and increasing the number of bits presented no great enhancement to the result. However, it is important to note that the quantification process induce a certain loss of
precision. The fixed point specification has to be applied to every operation and even every constant block in a model.

IV-THERMAL COMPUTATION RESULTS

In this study, investigations are done for the simplest case, only six temperature sensors (A,B,C,D,E and F) (Figure 2) in the form of two sensor cells and one single power heating, in order to validate with the 3-D FEM model. The simulations have been carried out for a one source placed at the junction level. As expected, the peak temperature profile is located at the centre of the heat source (figure 5). There is subsequent relaxation on temperature gradients through the structure leading to an essentially uniform temperature variation $\Delta T$. The sensor cells can be placed in any way out of the monitored area (different distance $H$ between cell 1 and cell 2), but in some cases adequate placement can simplify the thermal control unit design. In this part the results of thermal peaks can be very useful for indicating overheating situations and critical thermo-mechanical stress occurring in the device structure. Hence, table 1 display a comparison between the temperature peaks on surface with different implementation under the same conditions. Thus, the FEM results obtained are in full agreement with the GDS predictions.

<table>
<thead>
<tr>
<th>Detectors</th>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>V_A</td>
<td>68.655</td>
<td>98.655</td>
<td>75.324</td>
</tr>
<tr>
<td>V_B</td>
<td>70.248</td>
<td>100.248</td>
<td>76.324</td>
</tr>
<tr>
<td>V_C</td>
<td>71.325</td>
<td>101.325</td>
<td>77.055</td>
</tr>
<tr>
<td>V_E</td>
<td>69.365</td>
<td>99.365</td>
<td>75.603</td>
</tr>
<tr>
<td>V_F</td>
<td>70.325</td>
<td>100.325</td>
<td>76.540</td>
</tr>
<tr>
<td>V_G</td>
<td>72.318</td>
<td>102.318</td>
<td>78.649</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temperature peaks estimated on surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated by</td>
</tr>
<tr>
<td>FEM</td>
</tr>
<tr>
<td>SPTDA Float</td>
</tr>
<tr>
<td>SPTDA Fix</td>
</tr>
<tr>
<td>SPTDA Altera</td>
</tr>
</tbody>
</table>

Table 1 Temperature peaks comparisons

Fig. 4 Results generated using different fixed-point representations

Fig. 5 Temperature distribution according to A-C-axis [C°]

During implementation the fixed-point representation of the SPTDA algorithm was used. After implementation, the same input is directed to the algorithm simultaneously in Simulink® and on the FPGA board. The results are routed back to Simulink®, multiplexed, and projected on the same 2-D graph in order to compare the output. As many simulations and co-simulations have preceded the implementation, the result was expected to be the same as the comparison between the floating and the fixed point comparison. An optimal frequency close to 100 MHz has been achieved. Furthermore, the VHDL TB (Test Bench) was constructed and the force-file was used to stimulate
inputs and to compare with the algorithm predictions. Hence, the set of simulations revealed that the estimations generated by the SPTDA algorithm presented a great concordance with the predictions generated by the Finite Element Method (FEM) presented in [8].

V-THERMAL STRESS PREDICTION

If we know the peak temperature along with materials property and boundary conditions we can evaluate the thermal stress using the proposed approach. The algorithm presented herein might prove to be practical comparatively with thermal stress sensors methods especially for applications where the temperature has to be known only in a limited number of points, e.g. the determination of hot spot temperature or on-line temperature monitoring. The thermal stress (excluding intrinsic induced by packaging process) in thin films is given by the following expression:

\[ \sigma_{th} = \frac{E}{1 - \nu} \Delta \alpha \Delta T \]

Where \( E \) is the composite elastic constant for the different layers and \( \Delta \alpha \) is the difference in the coefficients of thermal expansion (CTE) between different level of packaging [9]. In [10] a method has been presented in details for calculation of the compressive stress in the silicon level given by:

\[ \sigma_{xx, max} = -9.63 \times (E \alpha_\text{Si} \Delta T_{in}) \]
\[ \sigma_{xx, max} = -9.63 \times (0.15 \times 2.8) \text{Si} \times (102.3 - 81.1) \]
\[ \sigma_{xx, max} = -8.574 \text{ MPa} \]

However, the induced compressive thermal stress will be combined to the intrinsic one due to the fabrication processes, and the stress due to the mechanical clamping structure.

VI-DISCUSSION

One of the important questions in the field of thermal issues of VLSI systems and microsystems is how to perform the thermal monitoring, in order to indicate the overheating situations, without complicated control circuits. Traditional approach consists of placement of many sensors everywhere on the chip, and then their output can be read simultaneously and compared with the reference voltage recognized as the overheating level. The idea of the proposed algorithm is to predict the local temperature and gradient along the given distance in a few places only on the monitored surface, and evaluate obtained information in order to predict the temperature of the heat source. Therefore, in the case of an SoC device, there is no place on the layout for the complicated unit performing computations, but there is also no need for it, as we only want to detect the overheating situations. These peaks are essentials during thermal die monitoring to avoid a critical induced thermo-mechanical stress. Moreover, in most cases, the overheating occurs only in one place.

VII-CONCLUSION

In this paper, a methodology to evaluate and predict a thermal peak of large VLSI circuit was presented. The important factors contributing to the device's thermal heating were characterized. The monitoring approach reported in this paper can be applied to predict the peak thermal stress of multilevel structures. Surface peaks thermal detection is necessary in modern VLSI circuits; their internal stress due to packaging combined with local self-heating becomes serious and may result in large performance variation, circuit malfunction and even chip cracking. In this paper GDS technique was used to develop SPTDA algorithm. Several approaches were implemented to achieve a better performance for the SPTDA operation. A new modular architecture that applied pipelining techniques and used some components in a time division fashion in order to reduce their number as much as possible, while enhancing the performance of the system was implemented. This was achieved using only 4 adders, 8 subtractors, 1 squarer, 4 dividers, and 6 multipliers. Regarding the area and speed achieved by our implementation, this SPTDA algorithm can be applied to surface peaks thermal stress detection performing a very accurate estimate in short time and using less area. Thereby, deploying the SPTDA in sensor network can be made feasible in the future.

References