Thermal Analysis of the Shared Register File in VLIW Architectures*

Anya Apavatjrut[1], José L. Ayala[2], Marisa López-Vallejo[2]

[1] Department of Telecommunication Services and Usage, INSA (Lyon, France)
E-mail: {anya.apavatjrut}@insa-lyon.fr
[2] Departamento de Ingeniería Electrónica, Universidad Politécnica de Madrid (Spain)
E-mail: {jayala, marisa}@die.upm.es

Abstract

This paper presents a thermal model used to visualize and analyze temperature evolution in the shared register files found on VLIW systems. The model allows the analysis of several factors that have strong impact on the heat transfer: layout topology, placement and access policy. The results obtained can further be used in the design of temperature- aware compilers and place&route tools.

Keywords: thermal model, register file, VLIW architecture

1. Introduction

Continuing advances in semiconductor technology have allowed dramatic performance gains for general-purpose microprocessors and embedded systems. These improvements are both due to increasing clock rates as well as to advanced support for exploiting instruction-level parallelism and memory locality using the additional transistors available in each process generation. However, as a negative consequence, this causes a significant increase in power dissipation, due to the fact that the dynamic power is proportional to both clock frequency and switching capacitance (which increases with the number of devices and components integrated in the chip). Thus, despite continuous attempts to reduce voltages and to design lower power circuits, power dissipation levels have steadily increased with every new generation of processor-based systems.

As technology scales down, higher power consumption and smaller chip area results in higher power density, which in turn leads to higher power temperature on the chip [Brooks (2001), Donald (2004)]. In fact, extrapolating the changes in microprocessor organization and the device miniaturization, power density near to 200W/cm^2

*This work is partially supported by the Spanish Government Research Grant TEC2006-00739

[SmartCooling] is projected for the near future. This requires extensive efforts on cooling techniques which have shown to be complex and expensive.

Early considerations at different design levels of these thermal effects can lead to more accurate design parameter estimation and faster design convergence [Huang (2004)], which is finally translated into better designs and reduced design time. Moreover, the analysis and management of every design parameter with a strong impact on the thermal behavior can provide an effective way to minimize the total temperature of the chip as well as the temperature gradient. Both metrics degrade the circuit performance and the reliability of the die and the package.

While hardware solutions to temperature management problems are very important, software can also play an important role because it determines the circuit components exercised during the execution and the period of time for which they are used. In particular, compilers and source code transformations determine the data and instruction access patterns of applications, what shapes the power density profile. Also, the topology of the hardware modules and the placement of these components determine the temperature behavior.

In this paper we present a complete parameterized thermal characterization of one of the hardware modules that can be found in a VLIW architecture, the shared register file. The experimental approach analyzes the effect of this device topology, as well as its placement in the chip layout, on the temperature behavior of the chip when different applications are run.

The contributions of this paper are:

1. Definition of a mathematical model to analyze the temperature behavior of the registers found inside the register file of a VLIW architecture. This model is integrated in a complete simulator of VLIW architectures in order to use the bus activities and register file accesses as input parameters for the model.

2. Analysis of the effect on the temperature behavior of different register file topologies, module placements and access policies.

This paper is composed as follows: Section 2 presents the previous relevant works in this topic, while our proposed methodology and thermal model are briefly explained in Section 3. Finally, Section 4 covers some preliminary results.

2. Related Work

In recent years there has been an increasing interest to provide a detailed die temperature distribution [Su (2003), Li (2004)]. In these works, the authors present different detailed full chip thermal models. All these models have characterized the temperature distribution across the silicon die with different accuracy, and they usually can be solved efficiently. However, it has been reported recently that the

results achieved by these models present inaccuracy problems [Huang (2005)]. All these works provide an analytical method for studying the thermal distribution in the die of high performance processors but they require the complete knowledge of the layout details. Moreover, these models are constrained to the processor layout and cannot be easily extended to different target architectures.

There are other approaches which relay on dynamic measures to characterize the thermal behavior of the chip [López-Buedo (2002)]. These techniques, opposite to ours, require not only the complete knowledge of the target architecture but also the capability to modify it. Some other approaches like [Julien (2003)] or [Senn (2004)] also propose techniques based on electrical measures to develop a power consumption model. However, these works have not dealt with the thermal behavior of the chip. Our previous research work [Ayala (2006), Méndez (2006)] has also targeted the thermal modeling of systems. This paper presents a different approach based on simulation that increases the granularity in the register file of the memory unit and explores several factors with impact on the temperature behavior, like the topology layout or placement.

Our work is based on the thermal model presented in [Paci (2006), Atienza (2006)]. We increase granuality of the previous model by extending its capability of thermal behaviour analysis inside the register file and define thermal behavior inside the register file by taking into account various placements, topologies and running benchmarks

3. Thermal Model

For the development of the thermal model, a well known analogy between the electrical circuits and the thermal sources is exploited. The silicon die and heat spreader are composed of elementary cells in a cubic shape. The temperature for every cell is computed using an RC model. The size of the cell trades-off the simulation speed with the thermal accuracy.

Each cell is associated with a thermal capacitance and five thermal resistances. Four resistances are used to model the horizontal thermal spreading, whereas the fifth is used to model the vertical thermal behavior. The thermal conductivity (horizontal and vertical)respectively, and then capacitance of each elementary cell are computed as follows:

$$G_{hor} = K_{G(Si/Cu)} \cdot \left(\frac{h \cdot w}{l} \right) \quad G_{ver} = K_{G(Si/Cu)} \cdot \left(\frac{l \cdot w}{h} \right)$$

$$C = K_{C(Si/Cu)} \cdot l \cdot h \cdot w$$

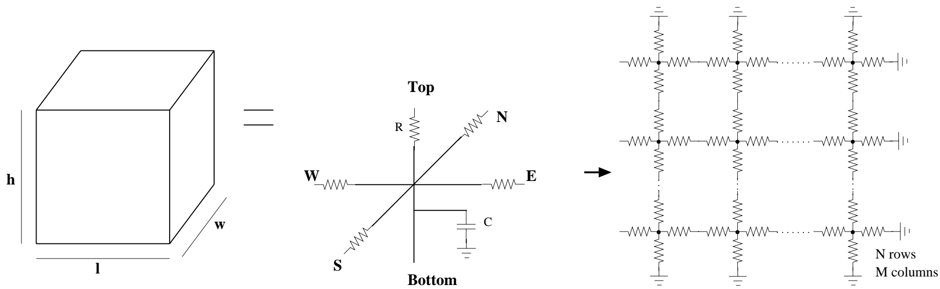

Figure 1. *Equivalent RC circuit of a cell.*

Figure 2. *Equivalent resistance circuit in a 2D map.*

where $K_{G(Si/Cu)}$ (thermal conductivity for silicon or copper), $K_{C(Si/Cu)}$ (thermal capacitance per volume unit for silicon or copper), l (cell length), w (cell width), h (cell height).

With this RC characterization, every cell is connected with the cells in the surroundings. The heat dissipation of each block is modeled as a source connected to the current node. A thermal circuit, which is similar to an electrical circuit, is created and can be solved by a node voltage analysis. As a result, the temperature of each block is obtained.

3.1 Register File Modeling

As was mentioned before, one of the goals of this work is to increase the model granularity by focusing the analysis on the temperature behavior of the registers inside the register file. To accomplish such goal, the register file is supposed to be represented as an $N \times M$ matrix and every register belongs to one of the elementary cells. Therefore, the thermal resistance and capacitance (Rcell and Ccell) for every elementary cell and every specific floorplan, have to be calculated.

Elementary resistances calculation

Since the total resistance and total capacitance of the device are known in advance, the register file can be decomposed into smaller units. Each unit is associated with its own resistance and capacitance as shown in Figure 1. This model will only consider those four resistances modeling the horizontal spreading of temperature. Therefore, the total resistance for a cell is

$$R_{cell} = R + \frac{R}{3} = \frac{4R}{3}$$

where R is provided by the inverse of G_{ver}.

Now, from Figure 2, the circuit can be decomposed in a matrix of N rows by M columns of registers, and the resistance of every row can be calculated as follows.

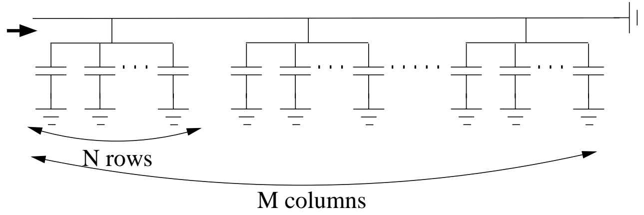

Figure 3. *Equivalent capacitance circuit.*

$$R_{cell,M} = \left[\left(R_{cell,M-1} + R \right) \| \frac{R}{2} \right] + R = \left[\frac{\left(R_{cell,M-1} + R \right) \cdot \frac{R}{2}}{R_{cell,M-1} + \frac{3R}{2}} \right] + R$$

Supposing that $R_{cell,M-1} = S_{M-1}R$ and $R_{cell,M} = S_M R$, then

$$S_M = \left\{ \left[\frac{S_{M-1} + 1}{S_{M-1} + \frac{3}{2}} \right] \cdot \frac{1}{2} \right\} + 1 \quad \text{with} \quad S_0 = 0$$

Considering that each row is parallel with the others, the total resistance for the device can be calculated by dividing by the number of rows.

$$R_{tot} = \frac{R_{cell,M}}{N} \quad \text{and} \quad R = \frac{NR_{tot}}{S_M}$$

Therefore, given the total resistance for the register file (value found in the thermal model from [Paci (2006)]), every register in the device can be characterized with its R_{cell}.

Elementary capacitances calculation

The total capacitance of the circuit can be calculated by considering each elementary capacitance to be parallel with the others (see Figure 3).

The total capacitance can be computed for N rows and M columns in parallel as

$$C_{tot} = C_{cell} \cdot N \cdot M \quad \text{and} \quad C = C_{cell} = \frac{C_{tot}}{N \cdot M}$$

Once the resistance and capacitance for each elementary cell are known, the size of the elementary cell can be calculated supposing that it is a quadratic cube by the expression

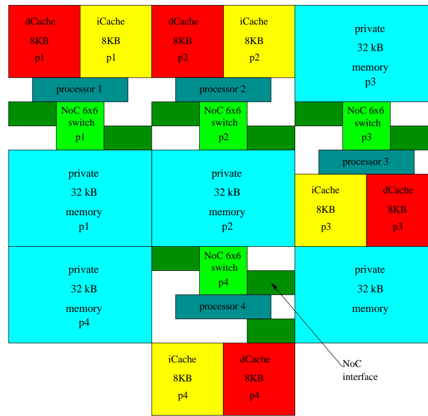

Figure 4. *MPSoC VLIW architecture.*

$$size = \frac{R}{K_{G(Si/Cu)}}$$

These last expressions are integrated in the VLIW simulator in order to retrieve the thermal behavior for every register in the register file for different placements and topologies.

From [Paci (2006)], we employ the same expression to compute the temperature evolution once the technology factors are calculated and the activities are obtained by the simulator.

$$T_c(n+1)\cdot 2^{36} = T_c(n) + ((cap \cdot EC \cdot 2^{62}) \cdot act) +$$

$$+((A\cdot 2^{26} - B\cdot 2^{26}\cdot(T_c(n))\cdot(T_n(n)\cdot 2^{36} - T_c(n)\cdot 2^{36})))/(2^{26})$$

where $T_c(n)$ is the temperature at step n, $T_c(n + 1)$ is the temperature at step n + 1, $T_c(n)$ is the neighbor cell temperature, cap × EC is the temperature difference due to the activities, act is the activity factor, A is the linear coefficient and B is the quadratic one.

4. Experimental Results

Once the thermal model for the register file has been developed, it has been integrated in the functional and thermal simulator of the VLIW system (see [De Beeck (2001)]).

In order to perform the thermal simulation of the system, a specific layout of the architecture has to be designed.

```
ddiiddiimmmm
ddiiddiimmmm
xppxxppxtttt
nssxnssxnssx
xssnxssnxssn
mmmmrrrrxppx
mmmmrrrriidd
ttttttttiidd
mmmmnssnmmmm
mmmmxssnmmmm
ttttxppxtttt
wwwwiiddwwww
wwwwiiddwwww
```

```
layer 1
p 1 1 1.0e-14
x 1 1 0.0
s 2 1 1.0e-14
i 2 1 1.0e-14
d 2 1 1.0e-14
n 1 1 1.0e-14
m 2 1 1.0e-14
t 1 1 1.0e-14
w 2 1 0.0
r 1 1 1.0e-14
```

Figure 5. *Configuration files.*

Figure 6. *Thermal map for different shapes : a) 1x64 b)2x32 c)4x16 d)8x8*

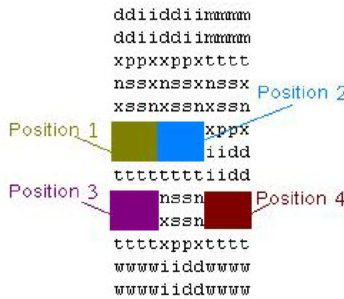

Figure 7. *Placements of the register file in the layout.*

The baseline architecture devised for the set of experiments resembles a common VLIW system with four processing cores, a shared memory subsystem, a shared register file and a communication network (see Figure 4). The layout of the system is configured in a text file where the placement and size of these modules is coded with letters (see Figure 5). One of the configuration files specifies the placement of the layout modules with a letter for every cell ('m' for the memory cell, 'r' for the register file cell, etc.), while the other configuration file provides the size of these cells.

During the simulation, the thermal coefficients and power coefficients for each cell are computed. Since every cell is surrounded by other cells, a heat distribution from hotter cells to colder cells takes place. At the end of the simulation, temperatures of all the cells are stored in a file.

A first set of simulations that have been carried out with the thermal model has focused on the analysis of the effect of the register file topology in the temperature behavior.

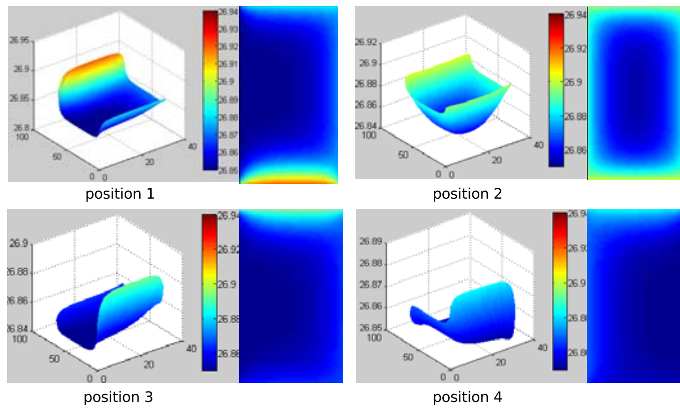

***Figure 8.** Thermal maps for different placements.*

Figure 6 shows the temperature map of a 64-position register file arranged in different configurations of rows and columns. A synthetic benchmark that performs intensive accesses to the device has been developed. As can be seen, when every register is equally accessed, the squared shape of the layout can mitigate the heat transfer between registers. In this case, the area in the center of the register file is found to be cooler than the borders. On the other hand, when a rectangular shape of the layout is used, the bottom registers suffer the heat transfer and they are found to be hotter.

A second set of experiments analyzes the effect of the placement of the memory unit on the thermal behavior. Therefore, four different placements in the layout have been defined (see Figure 7).

For every one of these placements, the thermal map of the register file is acquired supposing an homogeneous access pattern. Figure 8 shows these thermal maps.

As can be seen, when the register file is surrounded by hot devices as memories (position 2), or close to the processing unit (position 1), the temperature of the registers close to them is increased and thermal gradients between the hotter and cooler registers can appear destroying the silicon.

On the other hand, when the register file is placed near the border of the chip and the register file is surrounded by cooler devices (positions 3 and 4), the heat can be transferred to the outside of the system and the temperature is not increased. Moreover, the temperature map of the register file is homogeneous and the thermal gradients are reduced.

Before starting the simulation, the thermal parameters (the heat transfer coefficient and the thermal conductivity for silicon and copper) must be set. Also, the floorplan must be loaded.

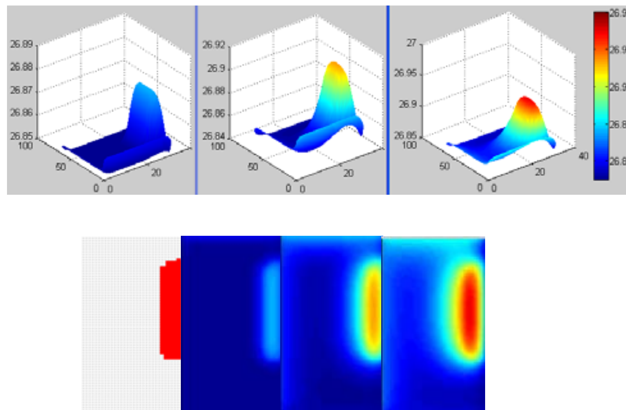

Figure 9. *Thermal map for a condensed access pattern.*

Finally, the last set of experiments analyzes the effect of the access pattern on the temperature of the register file. This analysis will allow to define temperature-aware access policies that reduce the temperature of the device as well as the power consumption [Atienza (2006)]. These experiments have been performed for every placement of the layout (positions 1, 2, 3 and 4) and three different access patterns (registers accessed from a bank placed on the right hand side of the register file, accessed registers randomly placed in several spots of the device and registers accessed in a homogeneous manner as a chess board).

The following graphs show the results for the three different accesses when the register file is placed in position 4.

Figure 9 shows the evolution in time of the thermal map for the register file when the registers are accessed from a bank located at the right hand side of the device. As can be seen, the bank where the registers are accessed from is increasingly heated as time advances. At the end, a large hot spot appears in the register file, what can severely damage the device.

A similar analysis has been done when the registers are randomly accessed from several spots in the device. The experimental results (not included due to space limitations) show how several hot spots appear on the register file surface increasing the probability of chip damage. Therefore, an access pattern that homogenizes the thermal map on the silicon surface must be found.

Figure 10 shows the evolution in time of the thermal map for the register file when the registers are accessed in a “chess board” manner. As can be seen, this access pattern homogenizes the temperature on the silicon because the accesses are distributed across a larger surface. Moreover, the probability of hotspots is minimized and the reliability of the system is not compromised.

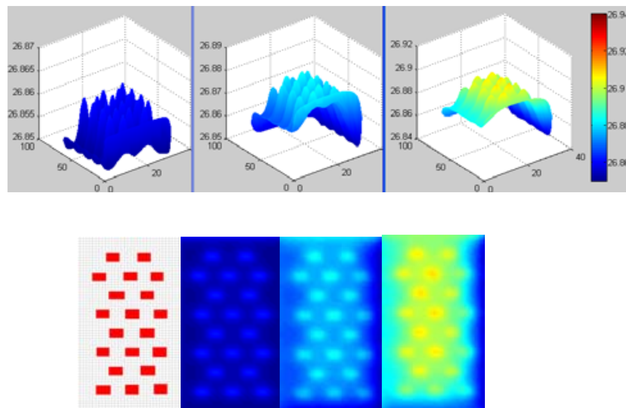

Figure 9. Thermal map for a « chess board » access pattern.

Therefore, it has been stated the strong impact of the register file topology, placement and access pattern on the thermal map of the device. These results open the way to new compilation and design strategies that optimize the temperature behavior on the modern VLIW systems.

5. Conclusions

There are several factors that have a strong impact on the thermal behavior of processor-based systems. Some of these factors, like the topology, placement or access policy, determine the thermal map of critic devices like the shared register file of VLIW systems.

This paper has presented an efficient analytical model to analyze the temperature evolution in the register file of the architecture, as well as to evaluate the factors that can modify this thermal map. The results obtained can be used in the design of temperature-aware compilers and place&route tools.

Acknowledgements

The authors would like to thank the contribution of the following researchers to this work, providing the simulation and thermal models used in our experiments: Giacomo Paci (Universita` di Bologna, Italy), Praveen Raghavan (IMEC, Belgium) and David Atienza (Universidad Complutense de Madrid, Spain).

References

http://www.hpl.hp.com/research/dca/smart cooling/

Atienza, D., Del Valle, P. G., Paci, G., Poletti, F. (2006), *A Fast HW/SW FPGA Based Thermal Emulation Framework for Multi-Processor System-on-Chip*. In Design Automation Conference.

Atienza, D., Raghavan, P., Ayala, J. L., de Micheli, G., Catthoor, F., Verkest, D., López-Vallejo, M. (2006). *Compiler-driven leakage energy reduction in banked register files*. In PATMOS.

Ayala, J. L., Méndez, C., López-Vallejo, M. (2006). *Analysis of the Thermal Impact of Source-Code Transformations in Embedded Processors*. In ICECS.

Brooks, D., Martonosi, M. (2001). *Dynamic thermal management for high-performance microprocessors*. In HPCA.

De Beeck, P., Barat, F., Jayapala, M., Lauwereins, R. (2001). *CRISP: A template for reconfigurable instruction set processors*. In FPL, pages 296–305.

Donald, J., Martonosi, M. (2004). *Temperature-Aware Design Issues for SMT and CMP Architectures*. In Workshop on Complexity-Effective Design.

Huang, W., Humenay, E., Skadron, K., Stan, M. (2005). *The need for a full-chip and package thermal model for thermally optimized IC designs*. In ISLPED.

Huang, W., Stan, M., Skadron, K. (2005). *Parameterized physical compact thermal modeling*. IEEE Trans. on Component Packaging and Manufacturing Technology, 28(4):615–622, December.

Huang, W., Stan, M., Skadron, K., Sankaranarayanan, K., Ghosh, S., Velusamy, S. (2004). *Compact thermal modeling for temperature-aware design*. In DAC.

Julien, N., Laurent, J., Senn, E., Martin, E. (2003). *Power consumption modeling and characterization of the TI C6201*. IEEE Micro, 23(5):40–49, September.

Li, P., Pileggi, P., Ashegi, M., Chandra, R. (2004). *Efficient full-chip thermal modeling and analysis*. In ICCAD.

López-Buedo, S., Garrido, J., Boemo, E. (2002). *Dynamically inserting, operating, and eliminating thermal sensors of FPGA-based systems*. IEEE Trans. On Components and Packaging Technologies, 25(4):561–566, December.

Méndez, C., Ayala, J. L., López-Vallejo, M. (2006). *Target Independent Thermal Modeling for Embedded Processors*. In IES.

Paci, P., Marchal, P., Polletti, P., Benini, L. (2006). *Exploring Temperature Aware Design in Low-Power MPSoCs*. In DATE.

Senn, E., Laurent, J., Julien, N., Martin, E. (2004). *Softexplorer: Estimation, characterization, and optimization of the power and energy consumption at the algorithmic level*. In PATMOS.

Su, H., Liu, F., Devgan, A., Acar, E., Nassif, S. (2003). *Full leakage estimation considering power supply and temperature variations*. In ISLPED.