Overheat protection circuit for high frequency processors

M. FRANKIEWICZ* and A. KOS

AGH University of Science and Technology, Department of Electronics, 30 A. Mickiewicza Ave., 30-059 Kraków, Poland

Abstract. The paper describes design and structure of the overheat protection circuit based on the PTAT sensors. The digital core of the system is driven by a 3-bit information generated by the structure. As a result, behaviour of the core differs for each temperature. The circuit was designed in LF CMOS 0.15 μm technology using full-custom technique. The presented paper focuses especially on the structure of the overheat protection circuit and simulations results of the functional blocks of the system. Layout and some parameters of the circuit are also considered.

Key words: PTAT, overheat, CMOS, VLSI, full-custom design.

1. Introduction

Modern technology of the integrated circuit fabrication is permanently scaling down [1]. As a consequence increasing power dissipation at the small area of the chip is often a cause of damage of the integrated circuit or at least errors in its behaviour. As an effect, thermal issues of the circuit work must be considered and real-time monitoring of the chip temperature is needed if there is a suspicion that the circuit can reach a high temperature range.

The information of the chip temperature can be created by many ways. One of possible solutions with many important advantages is to use PTAT (Proportional To Absolute Temperature) sensor. It has a linear temperature characteristic which is very suitable for further processing of the temperature-dependent signals. Very simple structure of the sensor allows the designer to create layout with very efficient utilisation of the chip area.

Next sections provide a concise description of overheat protection applications, more precise information of the PTAT sensor design and its usage in the presented overheat protection circuit. Finally, simulations of the prototype circuit and layout structure is presented.

2. Overheat protection applications

The problem of too big power dissipation resulting with overheating of the chip is present in almost every digital circuit, especially high frequency microprocessors. It has been already shown [2] that external heat removal methods (such as air cooling, microchannel cooling and heat sinks) are limited and do not fulfil requirements of modern integrated circuits. As an effect the thermal-aware design of integrated circuits is needed.

To protect the circuit from overheating the idea of Design for Thermal Testability (DfTT) has been introduced [3]. That means that a temperature sensor or even an array of sensors inside the package is needed. Based on that some sophisticated thermal monitoring systems have been designed. Paper [4] presents the idea of thermal Boundary Scan. In most cases the built-in temperature sensor monitors the temperature and generates a warning signal if temperature is near the dangerous range [5].

Another subject is how to manage the circuit behaviour when such a danger occurs. Dynamic losses are dependent mainly on the power supply value and a clock or control signal frequency according to (1)

$$P_{dyn} = C_{load} f V_{DD},\qquad(1)$$

where V_{DD} is supply voltage, f is signal frequency and C_{load} is capacitance of the load. As a result some systems limiting power consumption have been introduced including DVS (Dynamic Voltage Scaling), DCT (Dynamic Clock Throttling) and DFS (Dynamic Frequency Scaling) which can be combined [6, 7]. In standard applications all of mentioned methods sense the present temperature and after crossing the specified value they limit power consumption.

Some work has been concentrated on multi-core processors. For example [8] proposes the Thermal Spare Cores (TSC) method which saves one cooled core for a situation when another core is overheated. It seems to be an interesting approach but requires a large amount of additional circuit resources.

The described control methods which react on nearing to overheating need to be extended into more sophisticated systems which could provide the more efficient overheat protection. The next section provides a description of such a circuit.

3. Description of the chip

The presented overheat protection circuit consists of two main blocks: a set of PTAT sensors and the control block. An output signal of temperature sensors V_{PTAT} drives the control block. It is important to localize hot spots, so for further usage the maximum value of temperature should be taken. For test purposes the version of the layout with one sensor, located near

---

the point predicted as the hottest one, was prepared for fabrication. The temperature value is an important information for the core of the system so an output of the sensor block is additionally connected to the core. The control block uses information of the actual temperature to produce a 3-bit digital signal S_0–S_2 which controls the digital system core. The role of the signal S is to switch on or off chosen parts of the core. This arises from the necessity to reduce density of energy dissipated in the chip, which could be a reason of core overheating. The block diagram of the system is presented in Fig. 1.
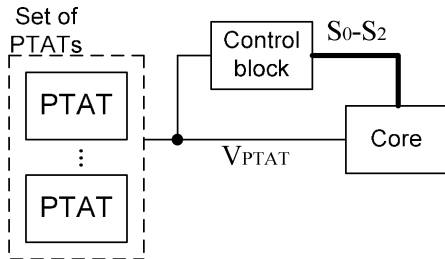

Fig. 1. Block diagram of the overheat protection circuit

The PTAT sensor task is to give the voltage proportional to the temperature of the chip. Its work is based on the phenomenon that difference between voltages of two diodes or bipolar transistors, which have different areas and conduct the same current, is proportional to the absolute temperature. The implemented structure, presented in Fig. 2, has been chosen from among others because its accuracy is sufficient for the specified application and presented circuit is easy to implement in CMOS process with planar transistors [9, 10]. Presented PTAT sensor consists of 5 MOS transistors, 2 bipolar transistors and 2 resistors.
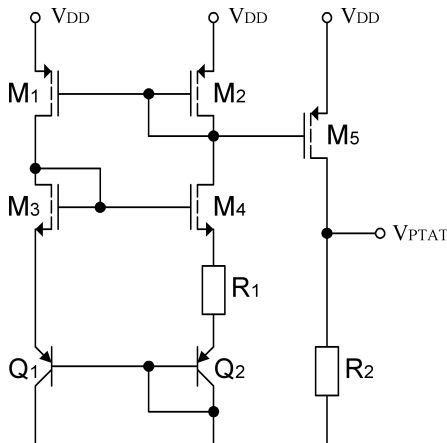

Fig. 2. Schematic diagram of the PTAT sensor

Some attention has to be given to output voltage of the sensor. Drain currents of transistors M_1 and M_2 are equal and they depend on present absolute temperature because of bipolar transistor included the lines. Current conducted by these transistor can be described by (2)

$$I_1 = I_2 = \frac{kT}{q}\frac{\ln(n)}{R_1},\qquad(2)$$

where k is the Boltzmann's constant, T is temperature, q is conducted charge and n is the ratio between emitter areas of Q_1 and Q_2 planar transistors. Transistors M_2 and M_5 create a current mirror. Consequently current conducted by the output transistor is equal to (3).

$$I_5 = \frac{W_5}{W_2}I_2 = \frac{W_5}{W_2}\frac{kT}{q}\frac{\ln(n)}{R_1}.\qquad(3)$$

After simple operations it can be calculated that the sensor produces the output voltage given by (4).

$$V_{\text{PTAT}} = I_5 R_2 = \frac{W_5}{W_2}\frac{kT}{q}\ln(n)\frac{R_2}{R_1}.\qquad(4)$$

A structure of the control block is similar to a flash Analogue-to-Digital Converter s(ADC) [11]. Temperature-dependent signal V_{PTAT} is compared with different reference voltages V_{REF}. This reference signals come from voltage dividers. Output signals from the comparators are decoded to 3-bit digital signal S_0–S_2. The schematic diagram of the control block is shown in Fig. 3.
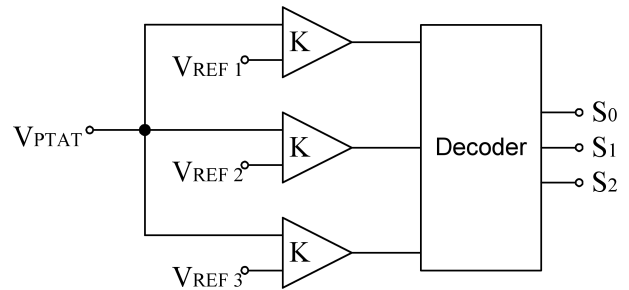

Fig. 3. Schematic diagram of the control block

The reason to form a decoder output into a 3-bit word is to make control of the core easier. The goal of signal S is to switch on and off different parts of the system core and each one of the S lines corresponds with another part of the core. The idea of switching off chosen blocks of the core leads to power consumption reduction. As an effect behaviour of the core differs for each temperature range. Structure of the decoder output word for each of the temperature ranges in presented case is shown in Table 1.

Table 1
Output word of the decoder in each temperature range

Temperature [°C]		Output word		
min	max	S0	S1	S2
–	60	0	1	1
60	75	1	0	1
75	90	1	1	0
90	–	1	1	1

A work of the overheat protection circuit should result in decreasing the temperature of the chip. Crossing consecutive temperature reference levels means that previous efforts were insufficient and another action is needed. At each temperature range power consumption of the chip is different. It has been decided to switch the whole core off when the level of the V_{PTAT} signal indicates that the chip temperature is higher than

acceptable and there is a risk of the error in a system work. If the work of the overheat protection circuit is efficient a chip temperature should decrease and the core returns to the previous state of work. That way the temperature feed-back loop is used to control the behaviour of the chip and secure the circuit from damage caused by the overheat.

Another issue is how to divide the core of the system into separate parts and how to ensure continuous correct work of the system while switching. The problem is quite wide so below only brief description and few suggestions are made. It is very important that the system core must maintain working properly without recognition in which working state it currently is. Obviously, after exceeding the highest temperature level that would cause serious failure, the system could be turned off and after cooling down to safe temperature range it could start working again correctly. Such solution is presented in this work. In case of lower temperature ranges the core division is more complex problem and there are few possible solutions. First of them is to design few blocks with similar functionality: more power-hungry but more efficient and low-power block with worse performance. Switching between such blocks would cause decrease of power consumption but is quite expensive in the case of chip area coverage and system complexity. Another possible solution for the energy-saving modes is to separate blocks which consume most of the energy and turn them off without any substitution. In such case tasks which should be executed by this blocks must be held in queue and wait for cooling and turning on a block referring to them. This type of division is quite problematic because it must be ensured that the rest of the system must be able to work correctly without disabled blocks. In other case when excluded parts are necessary the whole system must wait until they are awakened. The way of core division management depends on requirements of the specified system and in each case it must be considered by the system designers separately.

4. Simulation results

The presented circuit was designed in CMOS LF 0.15 μm technology with 1.8 V supply voltage and simulated in Cadence simulator. The layout of the chip was designed using full-custom technique and covers the area of 90×65 μm^2. Designed topography of the chip with marked main blocks of the circuit is shown in Fig. 4.
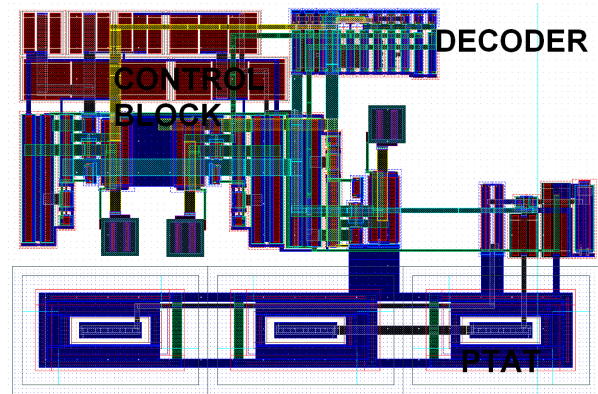

Fig. 4. Layout of the overheat protection circuit

The temperature sensor was designed using two PNP planar transistors with the emitters area multiplication factor $n = 2$. The material of the resistors has significant influence on the parameters of this sensor [9]. In this work resistors were built out of highly resistive polysilicon layer. The V_{PTAT} signal increases with a slope of about 8.42 mV/°C and ranges from 774 to 1532 mV at expected temperature range of 10 to 100°C. The temperature characteristic of the PTAT sensor is shown in Fig. 5. A very good linearity of the temperature sensor response is clearly visible.
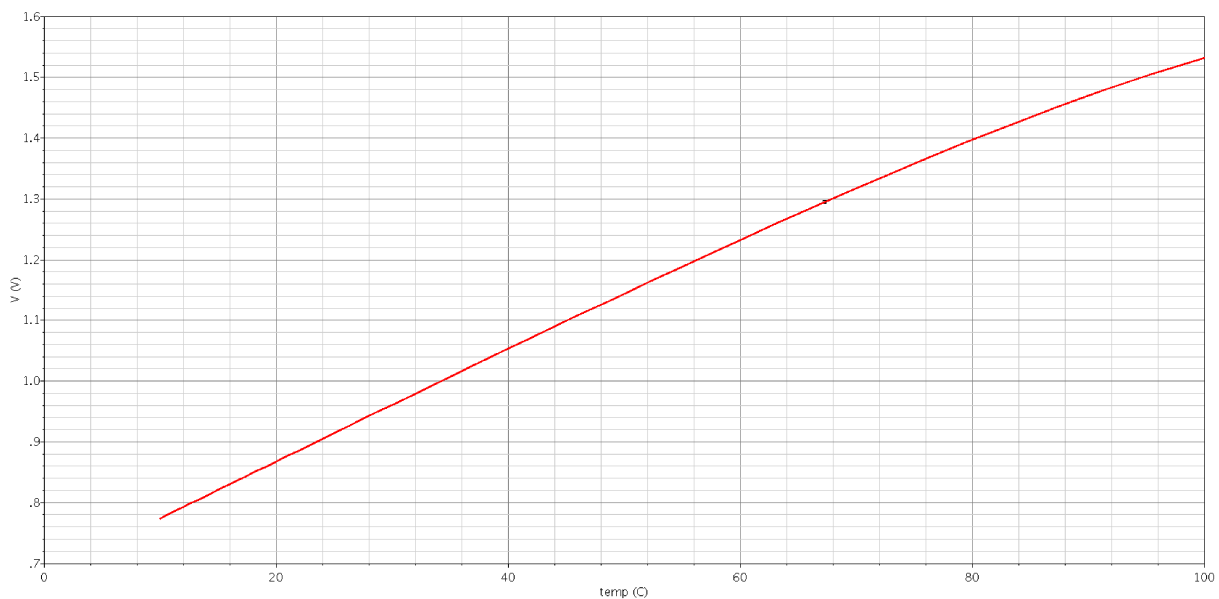

Fig. 5. Simulated characteristic of the PTAT temperature sensor
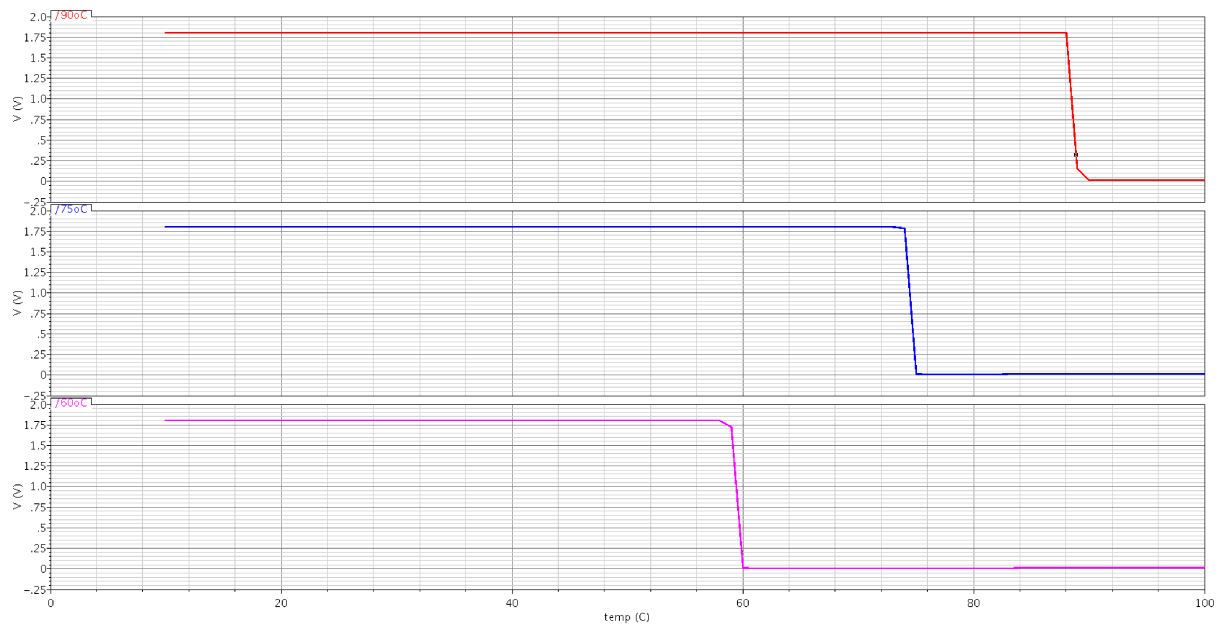
M. Frankiewicz and A. Kos

Fig. 6. Response of control block comparators in dependence of temperature. From top: output of comparator with V_{REF1} reference voltage (about 90°C), V_{REF2} reference voltage (about 75°C) and V_{REF3} reference voltage (about 60°C)
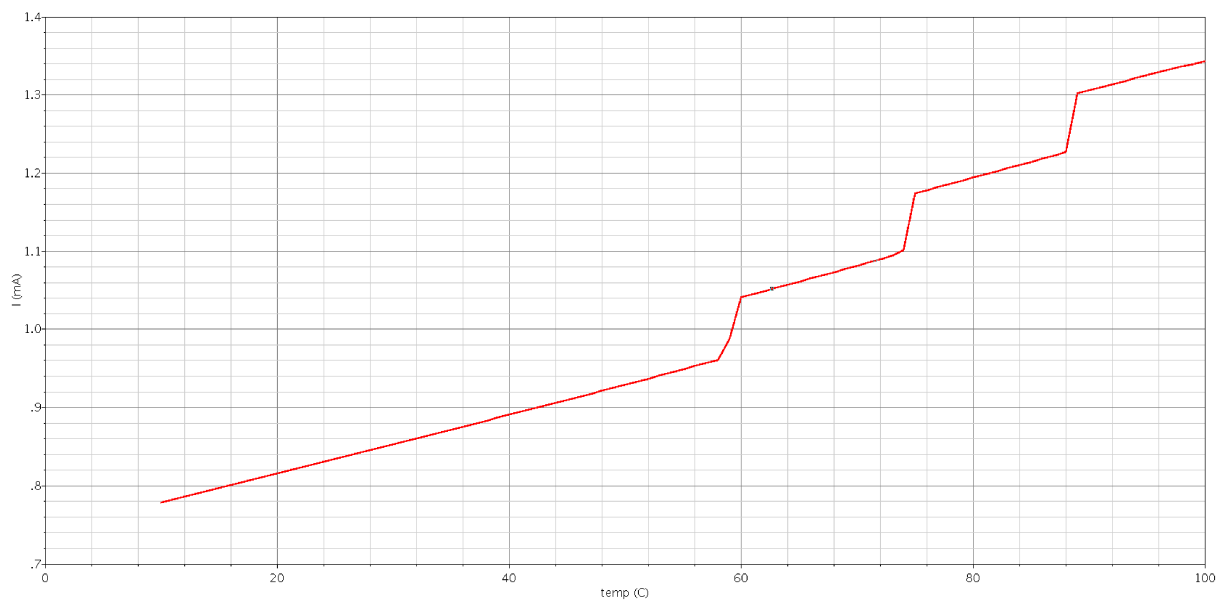

Fig. 7. Current consumed by the overheat protection circuit in dependence of temperature

The comparators in the control block were designed as two stage differential amplifiers with 10000 V/V amplification and 40 kHz of 3 dB-bandwidth. Reference voltages are produced by voltage dividers built of polysilicon resistors. The precision of the voltage division is ensured by a proper layout construction (the exact value of the resistors is not important, the good relation between resistances only matters). Decoder is a digital combinative circuit.

When temperature increases consecutive comparators change their outputs states to opposite ones. The temperature reference levels for each range are about 60, 75 and 90°C. Behaviour of the comparators outputs for different temperatures is shown in Fig. 6.

The circuit consumes about 1.4 to 2.4 mW (at 10 to 100°C) and power consumption is increasing with temperature what is show in Fig. 7. Total power consumption is a sum of two different parts. The linear increase is a result of fact that with higher temperature transistor M_5 in PTAT sensor conducts greater current. The three steps at about 60, 75 and 90°C result from switching the comparators in control block. An increase of the overheat protection circuit power consumption is not a problem because its work should result in greater power consumption decrease in the system core, so as an effect total system should be more energy-saving.

5. Conclusions

The paper presented simulation results and the structure of the overheat protection circuit already prepared for fabrication. The idea of the circuit is to use PTAT sensor to ensure safe work of the system core. The presented circuit makes state of chip work dependent on the present temperature. Introducing many temperature reference levels makes power dissipation control more fluent and as a result efficiency of chip work is better as well as overheating is less probable. Presented results show that the chosen structure successfully realizes its goal and can be competitive to previous applications. The designed circuit is planned to be fabricated and tested in a future chip.

The circuit presented in the paper is dedicated to a specified application but the idea with different reference temperatures and a proper decoder structure can be used in any real-time temperature monitoring system. Some further work with hysteresis is planned to improve the overheat protection circuit.

REFERENCES

[1] K. Gnidzińska, G. De Mey, and A. Napieralski, "Heat dissipation and temperature distribution in long interconnect lines", *Bull. Pol. Ac.: Tech.* 58 (1), 119–124 (2011).

[2] W. Huang, M.R. Stan, S. Gurumurthi, R.J. Ribando, and K. Skadron, "Interaction of scaling trends in processor architecture and cooling", *Proc. 26th IEEE Semiconductor Thermal Measurement and Management Symp. SEMI-THERM* 1, 198–204 (2010).

[3] V. Székely, C. Márta, Z. Kohári, and M. Rencz, "CMOS sensors for on-line thermal monitoring of VLSI circuits", *IEEE Trans. VLSI Systems* 5 (3), 270–276 (1997).

[4] V. Székely, M. Rencz, A. Páhi, and B. Courtois, "Thermal monitoring and testing of electronic systems", *IEEE Trans. Comp. and Pack. Techn.* 22 (2), 231–237.

[5] M. Szermer, M. Janicki, Z. Kulesza, and A. Napieralski, "Practical realization of PTAT sensor for ASIC overheat protection", *Proc. 15th Int. Workshop on Thermal Investigations of ICs and Systems THERMINIC'2009* 1, 80–83 (2009).

[6] S. Ankireddi and D. Copeland, "Thermo-mechanical reliability considerations with dynamic voltage/frequency scaling in microprocessor applications", *Proc. 25th IEEE Semiconductor Thermal Measurement and Management Symp. SEMI-THERM* 1, 134–138 (2009).

[7] A. Gołda and A. Kos, "Effective supervisors for predictive methods of dynamic power management", *Proc. 14th Int. Conf. Mixed Design of Integrated Circuits and Systems MIXDES'2007* 1, 381–386 (2007).

[8] M.A. Elsawaf, H.A. Hafmy, and A.L. Elshafei, "CPU dynamic thermal management via thermal spare cores", *Proc. 25th IEEE Semiconductor Thermal Measurement and Management Symp. SEMI-THERM* 1, 139–145 (2009).

[9] A. Gołda and A. Kos, "Analysis and design of PTAT temperature sensor in digital CMOS VLSI circuits", *Proc. 13th Int. Conf. Mixed Design of Integrated Circuits and Systems MIXDES'2006* 1, 415–420 (2006).

[10] A. Gołda and A. Kos, "Parameters identification of embedded PTAT temperature sensors for CMOS circuits", *Proc. 14th Int. Conf. Mixed Design of Integrated Circuits and Systems MIXDES'2007* 1, 392–395 (2007).

[11] M. Frankiewicz, P. Mroszczyk, A. Gołda, and A. Kos, "Asynchronous 4-bit CMOS flash analog-to-digital converter with over- and underflow detection system", *Proc. 16th Int. Conf. Mixed Design of Integrated Circuits and Systems MIXDES'2009* 1, 234–237 (2009).