Leakage-based On-Chip Thermal Sensor for CMOS Technology

Pablo Ituero José L. Ayala Marisa López-Vallejo
ETSI Telecomunicación (UPM)
Ciudad Universitaria s/n
28040 Madrid, Spain
{pituero,jayala,marisa}@die.upm.es

***Abstract*— Thermal characterization of ICs and on-chip temperature monitoring have become key tasks in electronic engineering. In this paper, we present the design of an on-chip CMOS temperature sensor based on the temperature dependent characteristics of the subthreshold current. The proposed sensor achieves high accuracy sensing (0.56^{o}C maximum error), wide temperature range (25-90^{o}C), and extremely low area (0.010 mm^2) and power overhead (18 μW). Our approach improves previous works on on-chip temperature sensors and is highly suitable for portable applications where temperature monitoring achieves great importance[1].**

I. Introduction

The increased chip densities and small features sizes achieved with the new sub-micron technologies have exploited the power density in the chip area and, consequently, the self-heating of the device is a very important factor. Self-heating introduces a branch of negative effects as electromigrations, poor signal integrity and chip damage. On-chip measurements of local temperature present an opportunity to incorporate temperature management techniques and performance optimization into the sensing chip (*self-consciousness*). Moreover, on-chip sensors can provide the needed feedback to verify the analytical models recently described [1] and incorporated in several CAD tools [2].

In CMOS technology, both MOS and bipolar transistors are commonly used to implement on-chip sensors, generating the signals for temperature sensors and voltage references [3]. In the case of MOS transistors, these signals are provided by the dependence of threshold voltage and mobility of carriers with temperature. The requirements for on-chip temperature sensing include not only high accuracy, but also compact layout area and very small power consumption to minimize the impact on the sensed magnitude.

External temperature sensors suffer a time-delay in the temperature reading due to the thermal constant from the external sensor to the integrated circuit [4]. Integrating the temperature sensor results in a lower cost solution that minimizes area penalty and provides instant information to enable thermal management.

In the last years, several researchers have proposed a variety of integrated on-chip temperature sensors. Chen et al. [5] designed a time-to-digital-converter-based CMOS sensor suitable for low-cost applications with limited measurement range. Pertijs et at. [6] developed a high-accuracy sensor by using dynamic element matching, a chopped current-gain bias circuit and a second-order sigma-delta ADC. The use of sigma-delta converters to implement temperature sensors robust from digital interference has been also proposed by Bakker and Huijsing [7]. Most of these approaches are based on traditional proportional-to-absolute (PTAT) principles and utilize bipolar transistors or lateral bipolars.

Some other groups have worked on the area of gradient detection instead of absolute temperature sensing. Shih et al. [8] describe one of these approaches with a large area implementation. The recent work by Zhai et al. [9] improves this idea with a tiny design suitable for the detection of temperature gradients.

In this work, we present the design of a CMOS temperature sensor based on the temperature dependent characteristics of the subthreshold current of MOS transistors. The increase in the subthreshold current achieved by the submicron technologies opens a very promising field in the development of new sensors. The proposed sensor is suitable for accurate on-chip temperature measures, is compatible with any CMOS technology and represents a low-area overhead. The most outstanding characteristics of the proposed sensor are:

- High accuracy and robustness in a wide temperature range.
- Suitable for sensing steady and quick thermal variations.
- Easy and cheap integration, the technology process is not compromised.
- Extremely reduced power and temperature overhead.

The paper is organized as follows. Section II makes a review of leakage currents in CMOS and their dependency on the temperature. In section III the thermal sensor is presented and two possible approaches to its implementation are described in section IV. The characterization of the circuit for both thermal-static and thermal-dynamic conditions is presented in section V. Finally section VI draws the main conclusions.

II. Temperature behavior of leakage currents

The continuous scaling of CMOS technologies has highlighted the importance of the leakage currents, those that flow even when the transistors are in an "off" condition.

[1]This work was funded by the CICYT project OPTIMA TEC2006-00739 of the Spanish Ministry of Science and Technology.

The leakage current consists of several components, from subthreshold conduction or reverse biased pn junction leakages to tunneling effects. However, the subthreshold conduction far exceeds the rest and characterizes the behaviour of the complete leakage current.

Subthreshold leakage is the weak inversion conduction current that flows between the source and the drain of a MOS transistor when gate voltage is below the threshold voltage, V_{TH} [10]. For the purposes of this work, the dependency of the subthreshold current on the device temperature is modeled as follows:

$$I_{DSUB}(T) = I_{S0}(T)e^{\frac{V_{GS}-V_{TH}(T)}{nkT/q}}(1 - e^{-\frac{V_{DS}}{kT/q}}) \quad (1)$$

In this equation kT/q is the thermal voltage, n is the transistor subthreshold swing coefficient, V_{GS} and V_{DS} denote the transistor gate-source and drain-source voltages respectively. $I_{S0}(T)$ is a technology parameter, dependent on the temperature given by:

$$I_{S0}(T) = \mu_0 C_{ox}\frac{W}{L}e^{1.8}(kT/q)^2 \quad (2)$$

where μ_0 is the zero bias carrier mobility, C_{OX} the gate oxide capacitance, W is the transistor width and L denotes the transistor effective channel length. Finally, in equation 3, $V_{TH}(T)$ denotes the transistor threshold voltage which can be expressed as follows as a function of the temperature:

$$V_{TH}(T) = V_{TH0} + (K_1 + \frac{K_2}{L} + K_3 V_{BS})(\frac{T}{T_0} - 1) \quad (3)$$

where V_{TH0} is the threshold voltage at a nominal temperature T_0, K_1 is the temperature coefficient of the threshold voltage, K_2 is the channel-length coefficient of the threshold voltage's temperature dependence, K_3 is the bulk-bias coefficient of the threshold voltage's temperature dependence, and V_{BS} is the transistor body-source voltage.

Previous equations show that the dependence of the subthreshold current on the temperature is complex and includes exponential and quadratic terms. However simulation results have proven that an exponential behavior provides an excellent fit for this dependency. Specifically, in the case when $V_{GS} = 0$, employed by this work, the subthreshold leakage can be modeled as follows:

$$I_{DSUB}(T) = K_A e^{\frac{-K_B}{T}}(1 - e^{-\frac{V_{DS}}{kT/q}}) \quad (4)$$

where K_A and K_B are empirical parameters. To summarize, given $V_{GS} = 0$, I_{DSUB} exhibits a positive exponential dependency on the temperature and its value decays also exponentially as V_{DS} becomes smaller.

III. Description of the Sensor

In order to take advantage of the variations of I_{DSUB} with the temperature, this work proposes to measure the time that I_{DSUB} takes to discharge a capacitance. This time will display the same thermal dependencies as the subthreshold current and will be easily quantified.

The structure of the sensor is displayed in figure 1. Similarly to what happens in dynamic gates, when the input receives a low-to-high transition and transistor M1 goes from an "on" to an "off" condition, capacitor C_L stores a charge that ideally would remain untouched, but that actually will gradually leak away due to leakage currents.

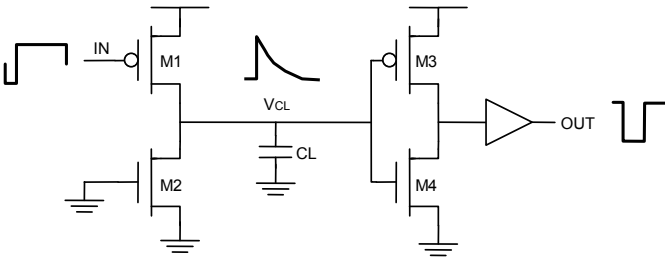

Fig. 1. Subthreshold current thermal sensor.

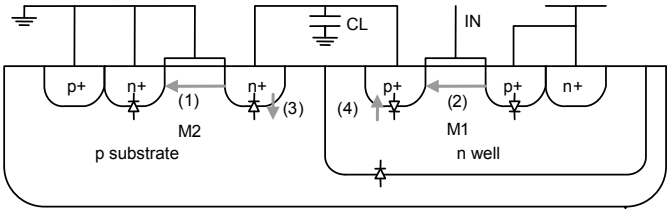

Fig. 2. Leakage current mechanisms in the thermal sensor.

Figure 2 shows the sources of leakage for transistors M1 and M2. Sources (1) and (2) are subthreshold leakages of M2 and M1, respectively. As mentioned before, these two components will dominate the behavior of the sensor. Sources (3) and (4) are reverse-biased diode leakages of M2 and M1 respectively. Sources (1) and (3) discharge C_L whereas sources (2) and (4) charge it. When C_L is charged, the transistor M2 drain-source voltage ($V_{DS,M2}$) equals Vdd and its subthreshold current ($I_{DSUB,M2}$) reaches a maximum. At the same time, $V_{DS,M1}$ and $I_{DSUB,M1}$ equal 0.

As the charge leaks away, $V_{DS,M2}$ decreases and so does $I_{DSUB,M2}$. In contrast, $V_{DS,M1}$ increases, which yields a rise in $I_{DSUB,M2}$. When both currents become equal, the voltage drop at C_L settles to an intermediate value determined by the ratio between the subthreshold current of the NMOS and PMOS devices, actually controlled by the sizing of the transistors. The sizing of the pull-down device is done in such a way that $I_{DSUB,M2} \gg I_{DSUB,M1}$ and the settling intermediate voltage is very close to 0V. In this way, this intermediate voltage is taken far from the switching threshold of the inverter M3-M4, moreover, when the measurement is captured, $I_{DSUB,M1}$ is still very small and can be neglected.

When the voltage at C_L falls below the switching threshold of the inverter M3-M4, the output produces a low-to-high transition. The time between the low-to-high transition at the input (IN signal in figure 1) and the low-to-high transition at the output (OUT signal in figure 1) is the measurement of the sensor. The magnitude of $I_{DSUB,M2}$ and its thermal dependencies will control this time interval. As a matter of fact, M2 can be considered the sensing device of the whole system. Figure 3 shows the functioning of the sensor. Note that the thermal dependency of the threshold voltages of M3

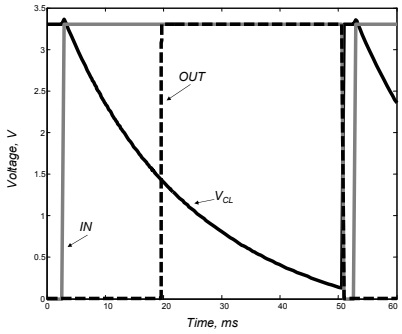

Fig. 3. Basic functioning of the thermal sensor.

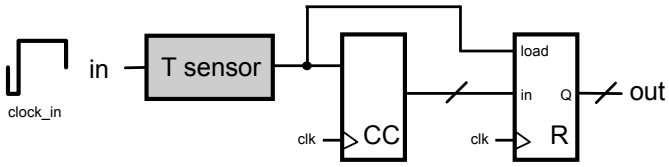

Fig. 4. Fixed implementation of the thermal sensor.

and M4 will impact the execution of the system and it will be integrated in the exponential response.

The robustness of the system relies on the stability of $I_{DSUB,M2}$ and the switching threshold of the inverter M3-M4; in other words, the sensibility of the sensor is restrained by the variations of these two parameters. In order to provide transistors M2, M3 and M4 with an improved stability, their transistor widths and channel lengths are higher than the minimum of the technology to assure that the manufacturing error is minimized.

IV. Interfacing

The output of the sensor yields an analog signal which is the duration of a pulse. This measurement must be converted into a digital form. In this section, two approaches to a digital interface for the sensor will be described.

A. Fixed approach

In the first approach, shown in figure 4, a clock with a very narrow low pulse —*clock_in*— drives the input. The measurement is performed by a cycle counter —*CC*— that measures the duration of the low pulse at the output of the sensor and stores the result in a register. Note that the loading rate of the register, i.e. the conversion rate of the sensor, is equal to the frequency of *clock_in*, thus constant. This frequency is bounded by the pulse width of the minimum temperature that the sensor needs to measure.

B. PWM approach

The second implementation is displayed in figure 5. In this case, the output of the sensor is stored in a D flip-flop and fed back to the input of the sensor one cycle later, so the capacitor is charged again one cycle after the switching threshold is achieved. A T flip-flop triggered by the output signal yields a clock signal whose high and low pulses last for the duration of the measurement. Again a cycle counter counts the number

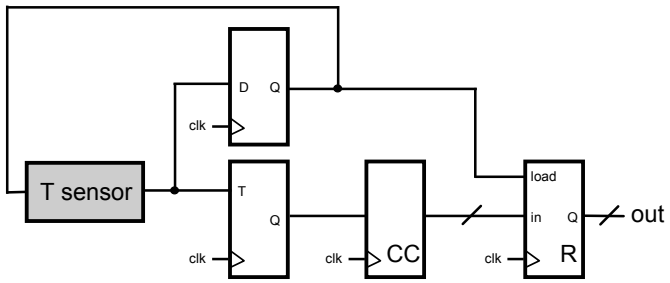

Fig. 5. PWM implementation of the thermal sensor.

of cycles between each transition and stores the result in a register. Now the register is uploaded at every transition of the sensor, therefore at a varying frequency that also depends on the temperature.

V. Characterization

The circuits proposed in this work have been tested using a 0.35μm technology from Austria MicroSystems^{TM}that employs a dual-poly quadruple-metal CMOS process. The experimental work and simulations have been carried out in the Cadence^{TM}environment. Both the fixed and the PWM implementation took up an area of 0.010 mm^2, which supposes a big improvement considering that the best predecessor, [5], achieved an area of 0.09 mm^2 for the same technology. As far as power consumption is concerned the fixed implementation, under a 3.3V power supply and a rate of 62.5 samples/s, the sensor attains 18.04μW. In the case of the PWM approach, since there is a variable conversion rate, the power consumption will also vary with the temperature; at a temperature of 60^oC the implementation consumes 86.30μW. Previous works do not provide information on power consumption at elevated conversion rates, so no comparison is possible.

A. Temperature-Static Behavior

Our first goal was to verify that the behavior of the sensor under static thermal conditions matched the exponential dependence on the temperature that theoretical equations foresee. The experiment consisted in measuring the pulse width at the output of the sensor at temperatures from 25^oC to 90^oC with no variation of the temperature for the duration of each measurement. Figure 6 displays the results employing a logarithmic scale for the pulse width. As shown, the tendency of the measurements corresponds very closely to the exponential dependency that the equations predicted. As a matter of fact, the data has been linearly fit in the logarithm domain with the least square method, so the average error is minimized and actually it is very close to 0^oC. The standard deviation of the error was found to be 0.33^oC and the maximum error 0.56^oC, these values approximate very closely all previous works and exceed the standard requirements of on-chip sensing.

B. Temperature-Dynamic Behavior

When the sensor measures a varying temperature, a distortion, produced by a finite conversion rate is produced and

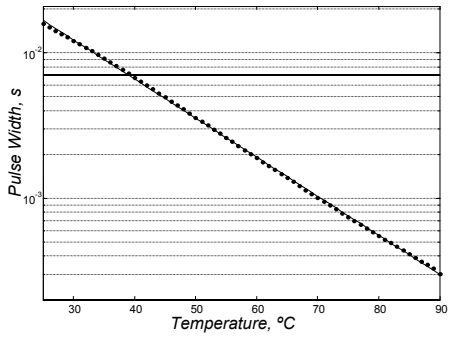

Fig. 6. Response of the sensor in the 25^{o}C to 90^{o}C range.

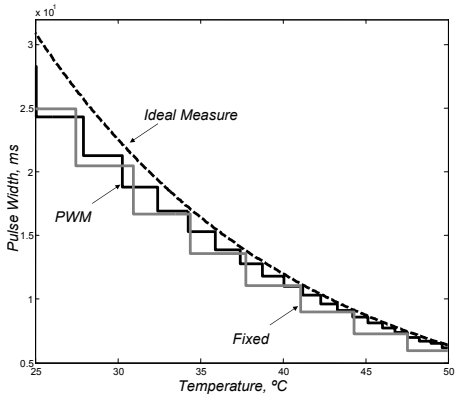

Fig. 7. Differences in response between the fixed and PWM approaches.

the result has to be analyzed to predict under- and overestimations. To evaluate the magnitude of this distortion the sensor was stressed with linearly varying temperatures from 90^{o}C to 25^{o}C at different rates. Since the maximum pulse width along this temperature interval is almost 16ms at 25^{o}C, temperature time-gradients above this value can be considered as quasi-static. Simulations aimed at temperature variation rates of 10ms/^{o}C and 1ms/^{o}C for both the PWM and the fixed approaches. Figure 7 illustrates the differences between both implementations, note that while the PWM implementation approximates the ideal response as the temperature rises, the fixed implementation maintains a constant rate and therefore a constant error along the temperature range. To the best of our knowledge no previous work has analyzed the effect of changing temperatures before. Our work assures the stability of the sensor output under different working temperatures.

Table I shows the means and standard deviations of the errors for the fixed implementation. In this case, the average error remains constant for the whole range of temperatures. The inaccuracy of the sensor grows rapidly beyond the 1ms/^{o}C decade because of the limited conversion rate. The results for the PWM implementation are shown in table II. In this case, apart from the statistics of the whole temperature range, the table shows the values for four sections of the range. As shown, the average error decreases as the temperature rises, which is caused by the growth in the conversion rate. In comparison with the fixed implementation, PWM exhibits a superior performance for all the temperature range and at every temperature rate, albeit the power consumption also increases.

TABLE I

ERROR STATISTICS FOR THE FIXED APPROACH.

10ms/^{o}C		1ms/^{o}C	
Mean (^{o}C)	SD (^{o}C)	Mean (^{o}C)	SD (^{o}C)
1.01	0.56	9.69	5.03

TABLE II

ERROR STATISTICS FOR THE PWM.

	10ms/^{o}C		1ms/^{o}C	
	Mean (^{o}C)	SD (^{o}C)	Mean (^{o}C)	SD (^{o}C)
Full Range	0.45	0.28	3.46	3.21
25^{o}C-45^{o}C	0.73	0.58	7.35	3.07
45^{o}C-60^{o}C	0.48	0.15	2.97	0.89
60^{o}C-75^{o}C	0.70	0.07	1.73	0.43
75^{o}C-90^{o}C	0.30	0.19	0.63	0.39

Thus, the selection of the best implementation must consider a trade-off between accuracy and power overhead.

VI. CONCLUSIONS

The design of a CMOS on-chip temperature sensor based on the thermal characteristics of the leakage currents has been presented. The device is compatible with any CMOS technology and features high accuracy and robustness under typical thermal conditions of high-end processors, furthermore it exhibits a suitable performance for quick temperature variations. The sensor overcomes all previous works as far as area (0.010 mm^2) and power (18 μW for a fixed rate of 62.5 samples/s) are concerned.

REFERENCES

[1] H. Pape et al, "Thermal Transient Modeling and Experimental Validation in the European Project PROFIT," *IEEE Transactions on Components and Packaging Technologies*, vol. 27, no. 3, pp. 530–538, September 2004.

[2] W. Batty et al, "Electro-Thermal CAD of Power Devices and Circuits with Fully Physical Time-Dependent Compact Thermal Modelling of Complex Non Linear 3-Dimensional Systems," *IEEE Transactions on Components and Packaging Technologies*, vol. 24, no. 4, pp. 566–590, December 2001.

[3] G. C. M. Meijer, G. Wang, and F. Fruett, "Temperature sensors and voltage references implemented in CMOS technology," *IEEE Sensors Journal*, vol. 1, no. 3, pp. 225–234, October 2001.

[4] H. Sanchez, R. Philip, J. Alvarez, and G. Gerosa, "A CMOS Temperature Sensor for PowerPC^{TM}RISC Microprocessors," in *Symposium on VLSI Circuits*, 1997, pp. 13–14.

[5] P. Chen, C.-C. Chen, C.-C. Tsa, and W.-F. Lu, "A Time-to-Digital-Converter-Based CMOS Smart Temperature Sensor," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 8, pp. 1642–1648, August 2005.

[6] M. A. P. Pertijs, K. A. A. Makinwa, and J. H. Huijsing, "A CMOS Smart Temperature Sensor with a 3σ Inaccuracy of $\pm$0.1 ^{o}C From -55 ^{o}C to 125 ^{o}C," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2805–2815, December 2005.

[7] A. Bakker and J. H. Huijsing, "Micropower CMOS Temperature Sensor with Digital Output," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 7, pp. 933–937, July 1996.

[8] Y.-H. Shih, S.-R. Lin, T.-M. Wang, and J.-G. Hwu, "High Sensitive and Wide Detecting Range MOS Tunneling Temperature Sensors for On-Chip Temperature Detection," *IEEE Transactions on Electron Devices*, vol. 51, no. 9, pp. 1514–1521, September 2004.

[9] Y. Zhai, S. B. Prakash, M. H. Cohen, and P. A. Abshire, "Detection of On-Chip Temperature Gradient Using a 1.5V Low Power CMOS Temperature Sensor," in *IEEE International Symposium on Circuits and Systems*, 2006, pp. 1171–1174.

[10] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI Devices*. Cambridge University Press, 1998, pp. 120–128.