Delay-line based Temperature Sensors for On-chip Thermal Management

Shuang Xie and Wai Tung Ng

The Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto, 10 King's College Road
Toronto, ON, Canada, M5S 3G4
shuang.xie@utoronto.ca; ngwt@vrg.utoronto.ca

Abstract

Integrated digital temperature sensors facilitate advanced thermal and power management. This paper reviews the integrated delay-line based temperature sensors, in terms of operating principle, the state-of-the-art power and area optimization and calibration methods. A self-calibration approach recently introduced will also be discussed in detail. This self-calibration method allows the automatic elimination of process variations and mismatches without the need for individual preprocess trimming as needed in traditional approaches. Measurement results for a 65nm CMOS delay-line based temperature sensor confirms an energy per conversion of 0.02 nJ with a resolution of 0.5 °C between 20 to 80 °C with a maximum error of ±2.0 °C. The active area of the temperature sensor is only 0.002 mm^2.

1. Introduction

The power density in modern microprocessor chips continue to elevate as VLSI technology scales. To avoid heat dissipation at certain locations from reaching destructive conditions, temperature is sensed and the system operation is adjusted accordingly [1]. As a result, accurate sensing of the chip temperature at critical locations becomes necessary. For instance, AMD's Opteron microprocessor utilizes 38 temperature sensors as part of its thermal management system [2].

This paper is organized as follows. Section 2 covers the operating principles of the state-of-the-art integrated delay-line based temperature sensors. The power and area optimization methods, along with calibration methods, are addressed in Section 3. In particular, delay-line based temperature sensors with self-calibration are discussed in detail. This is followed by the conclusions and future trends in Section 4.

2. Operating Principles

2.1 Sub-categories in smart temperature sensor

Smart temperature sensors, or digital temperature sensors, can be divided into two categories: bandgap voltage based and delay-line based. The bandgap temperature sensor relies on a pair of BJT devices with different bias currents to generate a PTAT (proportional to absolute temperature) voltage [3]. In modern CMOS technologies, the BJTs are usually area consuming lateral PNPs. However, in applications where a large amount of temperature sensors are needed (as shown in Fig. 1) and accuracy (e.g. 3-5 °C) [1] is not the primary concern, the delay-line based temperature sensors consume much less area and lower power [4]-[14].

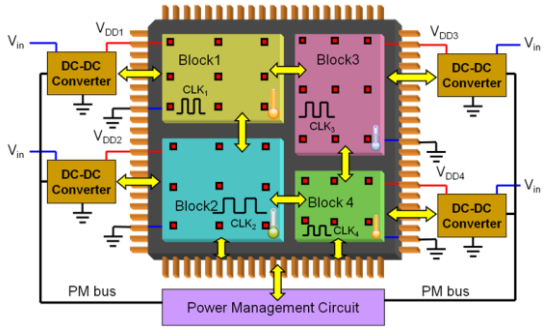

Figure 1. A conceptual diagram, showing the incorporation of an array of temperature sensors as part of the on-chip VLSI power management.

There are three sub categories of delay-line based temperature sensors: 1) time domain, 2) frequency domain, 3) voltage domain. All these delay-line based temperature sensors contain two main functional blocks to perform temperature sensing and quantization. For the time domain design, there are two delay lines in the sensor, one for temperature to time conversion and another for time to digital conversion. A counter is often used to facilitate the quantization. In the frequency domain design, ring oscillator converts temperature into frequency and the frequency is quantized by a counter. In the voltage domain design, a voltage controlled delay line is used for the voltage to digital conversion. The temperature sensing device is usually MOSFETs operating in weak inversion.

2.2 Thermal sensing circuitries

All delay-line based temperature sensors rely on two

temperature correlated parameters for thermal sensing: surface carrier mobility μ and threshold voltage V_{TH}. The propagation delay of a logic inverter cell (operating in strong inversion region) exhibits positive temperature coefficient [4]:

$$T_p = \frac{2(L/W)C_L V_{TH}}{\mu C_{ox}(V_{DD}-V_{TH})^2} + \frac{(L/W)C_L}{\mu C_{ox}(V_{DD}-V_{TH})}\ln\left(\frac{1.5V_{DD}-2V_{TH}}{0.5V_{DD}}\right) \quad (1)$$

$$\mu = \mu_0\left(\frac{T}{T_0}\right)^{-\alpha_u}, \alpha_u = 1\sim 3;\ V_{TH}(T) = V_{TH}(T_o) - \alpha_v(T-T_0)$$

where T_p is the propagation delay of a single cell inverter (charging or discharging). This delay is dominated by the negative temperature coefficient of mobility, μ [5]. Equation (1) is the basis for time domain sensors [6]-[8] and frequency domain sensors [9]. However, when operating in the weak inversion region (subthresold), the inverter's propagation delay exhibits a negative temperature coefficient that is dominated by the threshold voltage [10], [11]. Measured delays in TSMC's 65nm CMOS technology under strong and weak inversions are as shown in Fig. 2 and Fig. 3, respectively. The MOSFETs can be used for creating voltages with positive and negative temperature coefficients, depending on bias conditions [12].

2.3 Quantization

In a time domain temperature sensor, the quantization is performed by a time to digital converter (TDC). In [5] and [8], the TDC is an inhomogeneous cyclic delay line that shrinks the input pulse by a fixed amount of time after each cycle until it diminishes completely. A counter determines the number of pulses and generates digital representations of the temperature. In [13], the TDC is a time comparison circuitry using successive approximation algorithm. In [7], the TDC is a DLL (delay locked loop). Frequency domain temperature sensor combines the temperature sensing and quantization using a ring oscillator [9]-[11]. The oscillation frequency is temperature dependent and is inversely proportional to the propagation delay of the delay cells [10], [11]. The block diagram of such structure is as shown in Fig. 4.

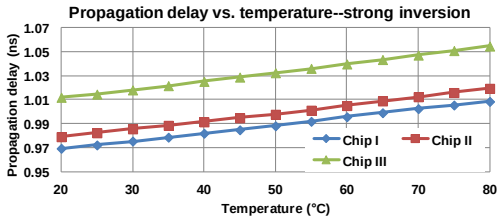

Figure 2. Measured propgation delays from mulitple 64 inverters delay lines showing positive temperature coefficient when operating in strong inversion mode.

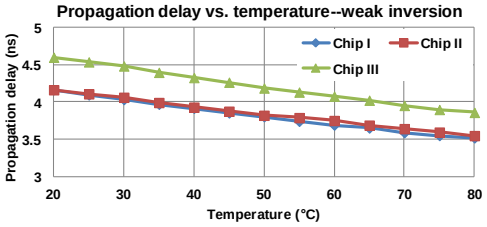

Figure 3. Measured propgation delays from mulitple 64 inverters delay lines showing negative temperature coefficient when operating in weak inversion mode.

3. State-of-the art technologies

3.1 Power and area saving methods

Ring oscillator and cyclic delay lines can reduce chip area at the expense of conversion time. Majority of delay-line based temperatur sensors operate at 1kHz or slower. To achieve higher resolution, the length of the temperature-to-time conversion delay line has to be increased. This length requirement for the delay line can be eliminated by the used of a cyclic temperature dependent delay line [6], [8]. The inverter chain's delay pulse width is multiplied by the number of cycles. As shown in Fig. 5, when "Enable" rises to high, the delay line starts to oscillate. The counter divides this oscillation pulse. When the number of pulses reaches a specified number, the counter is reset and its output pulse is passed to the XOR gate with the "Enable" signal to generate the multiplied pulse. In a ring-oscillator based temperature sensor, the resolution increases proportionally with the conversion time without the need to increase the length of the inverter chain. However, one disadvantage of the ring-oscillator based temperature

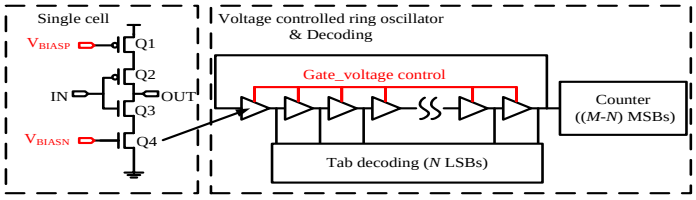

Figure 4. The block diagram of a frequency domain temperature sensor, showing the voltage controlled delay cell (left); temperature sensing and quantizing ring oscillator (right), and power saving method with counter and tab decoding [11].

TABLE 1. SPECIFICATIONS OF THE STATE-OF-THE-ART TEMPERATURE SENSORS

Ref	Technology	Architecture	Area	Power	Temperature Range	Resolution	Calibration Method	Accuracy	Energy per conversion
[3]	65nm	Bandgap	0.1 mm^2	10 µW	-70 ~125 ºC	0.03 ºC	Two-point	0.2 ºC	4.5 µJ
[6]	0.22µm	Delay line	N/A	175 µW	0 ~ 100 ºC	0.133 ºC	One-point	±0.7 ºC	175 nJ
[7]	0.13µm	Delay line	0.12 mm^2	1.2 mW	0 ~ 100 ºC	0.66 ºC	One-point	±2.3 ºC	0.24 µJ
[8]	65nm	Delay line	0.01 mm^2	150 µW	0 ~ 60 ºC	0.139 ºC	Auto	±5.1 ºC	15 nJ
[11]	65nm	Ring oscillator	0.002 mm^2	60 µW	20 ~80 ºC	0.5 ºC	Self-cal	±2 ºC	0.02 nJ
[14]	65nm	Ring oscillator	N/A	120 µW	20 ~80 ºC	0.16 ºC	One-point	±1.5 ºC	115 nJ

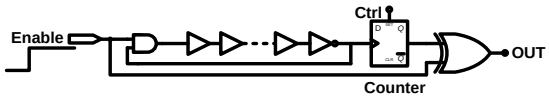

Figure 5. A cyclic delay-line can be used as a temperature sensing element for reducing the length of the invert chain.

sensor is the fact that the power consumed by the counter increases linearily with the oscillation frequency. To reduce this power, a tab decoding method is introduced in [11], as shown in Fig. 4. The total dynamic power and energy per conversion consumed by the ring oscillator and the counter can be expressed as:

$$\begin{aligned} P_{dyn} &= P_{dyn\_ringoscillator} + P_{dyn\_counter} \\ &= V_{DD}\cdot i_{ave} + V_{DD}\cdot i_{ave}\cdot\frac{C_{counter}}{C_{inverter}}\left(\sum_{x=0}^{M-N}\frac{1}{2^x}\right) \\ E &= P_{dyn}\cdot\frac{1}{f_s} = \frac{D_{coarse}}{2^N}\cdot V_{DD}^2\cdot\left(C_{inverter} + \frac{C_{counter}}{N_{stage}}\left(\sum_{x=0}^{M-N}\frac{1}{2^x}\right)\right) \end{aligned} \quad (2)$$

where V_{DD} is the supply voltage; i_{ave} and $C_{inverter}$ are the average discharging or charging current and the gate node capacitance in a single inverter; N_{stage} is the number of inverter cells in the ring oscillator; $C_{counter}$ is counter's LSB register's gate capacitance and D_{coarse} is the digital output of the sensor. Equation (2) shows that larger gate node capacitance will result in lower power consumption at the expense of increased energy per conversion, as demonstrated in [10]. At the same time, increasing N_{stage} will reduce the dynamic power consumed by the counter without any effect on the energy per conversion. On the other hand, increasing the decoder's number of bits, N reduces both energy and power, as a smaller counter can be used while maintaining the same resolution [11]. This method reduces the dynamic energy by a factor of more than 2^N.

3.2 Calibration methods

Fig. 2 and Fig. 3 show the process variations between measured digital outputs. These variations can be observed not only between chips but also on the same chip (see Fig. 6). To establish the direct relationship between digital outputs and temperature, the curves, as shown in Fig. 2 and Fig. 6, have to be calibrated at least at two different temperature points [5], to determine gain and offset for each individual sensors. However, if a large number of on-chip temperature sensors are needed in a microprocessor, it is impossible to perform two-point calibration individually. Therefore, an automatic one-point calibration method is proposed in [7]. It is based on the fact that the inverter delay varies with both temperature and process, and the two effects can be separated as [7]:

$$D(T,P) = T^{-\alpha_\mu}\cdot G(P) \quad (3)$$

where P denotes various process variations in $G(P)$. After eliminating $G(P)$, $T^{-\alpha_\mu}$ is the only temprature dependent term.

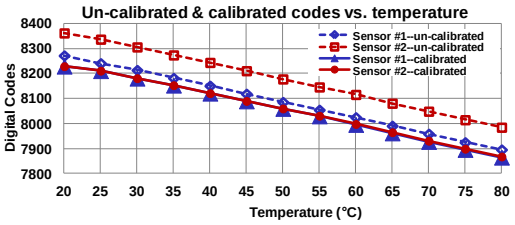

Figure 6. Measured results from two identical all digital temperature sensors programmed at different locations on a Cyclone III FPGA chip. The dash lines are before one-point calibration and the solid ones are after calibration [14].

The calibration process is described in detail in [7], and is briefly repeated here. At a particular calibration temperature (e.g. 50 °C), all the temperature dependent delay lines are adjusted to have the same delays. This is accomplished by comparing it with a reference (temperature independent) DLL with a fixed resolution. During measurement mode, the pulse width of the temperature dependent line is measured using the DLL, and digital reprentations of the temperature is generated.

The one-point calibration method proposed in [6] tries to make all the sensors' digital output codes to be the same at 50 °C, by adjusting the number of cycles in the temperature dependent cyclic delay lines using an off-chip time-domain phase detection circuit.

The one-point calibration approaches in [6] and [7] remove process variations between sensor outputs, as shown in Fig. 6. However, to establish the relationship between these calibrated uniform codes and temperature, they have to be followed by a linear (two-point calibration), or second order or third order curve fitting, as done in [6], [7]. The curve fitting is needed only once for each batch of sensors from the same technology.

An automatically calibration method proposed in [8] correlates measurement results with simulation results in the TT, FF and SS corners. The method's accuracy is dependent on the validity of the SPICE model. A self-calibration method is proposed in [11], and it relies on the positive and negative temperature coefficients of two ring oscillators, as their inverters work in weak and strong inversion separately. This is demonstrated by the measurement results in Fig. 2 and Fig. 3. As shown in Fig. 7, the calibration procedure involves seeking the proper gate voltage for driving the auxiliary line into PTAT without any offset. The output codes of the auxiliary line are used to help calculate the main oscillator's gain and offset. Measurement results using this self-calibration method is as shown in Fig. 8 [11]. A comparison chart on the state-of-the-art delay-line based temperature sensor is given in Table 1.

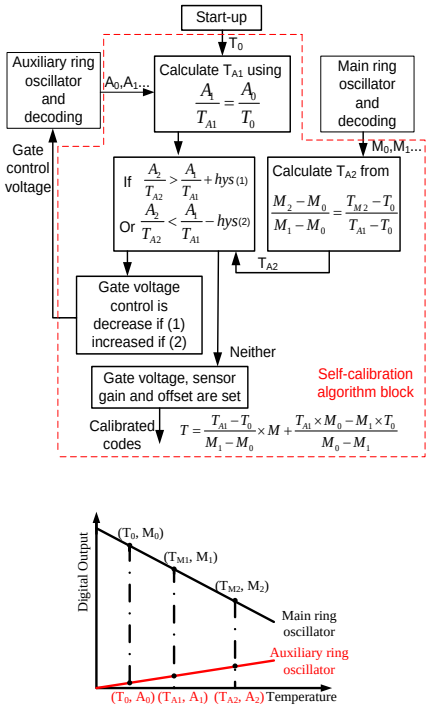

Figure 7. The self-calibration method as described in [11].

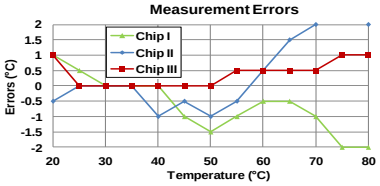

Figure 8. Measured temperature errors as seen in [11].

4. Conclusions

Distributed integrated digital temperature sensors are needed in microprocessor chips to implement thermal and power management. Self-calibration is an essential feature to facilitate mass production. Future scaling will lead to more mismatches and process variations. The all-digital temperature sensor will become more favorable as it is easier to synthesize and port to future technologies.

Acknowledgments

This work was supported in part by the China Scholarship Council (File No. 2009102021), AMD Canada, and Natural Science and Engineering Research Council of Canada. We would like to acknowledge Canadian Microsystems Corp. for facilitating the IC fabrication and design support.

References

[1] Y. William Li and Hasnain Lakdawala, in Proc. IEEE CICC, p.1(2011).

[2] Y. Zhang and Ankur Srivastava, IEEE DAC Conf., p.472(2009).

[3] F. Sebastiano, L.J. Breems, K.A.A. Makinwa, S. Drago, D.M.W. Leenaerts, and B. Nauta, IEEE J. Solid-State Circuits, p.1(2010).

[4] M.I. Elmasry, Digital MOS integrated circuits, p.1(1981).

[5] P. Chen and Shen-Iuan Liu, Custom Integrated Circuits, Proceedings of the IEEE, p. 605(1999).

[6] P. Chen, S.C. Chen, Y. S. Shen, and Y. J. Peng, IEEE Trans. Circuits Syst. I, p. 913 (2011).

[7] K. Woo, S. Meninger, T. Xanthopoulos, E. Crain, D. Ha, and D. Ham, in Proc. IEEE ISSCC Dig, p.68(2009).

[8] C.C. Chung and C.R. Yang, IEEE Trans. Circuits Syst. II, p.105(2011).

[9] C.K. Kim, B.S. Kong, C.G. Lee, and Y.H. Jun, IEEE ISCAS Conf., p. 3094(2008).

[10] S. Park, C.Min, and S.H. Cho, IEEE ISCAS Conf., p.1153 (2009).

[11] S. Xie and W.T. Ng, "A 0.02 nJ self-calibrated 65nm CMOS delay-line temperature sensor," to be published in Proc. IEEE ISCAS, May. 2012.

[12] K. Law, A. Bermak, and H.C. Luong, IEEE J. Solid-State Circuits, p. 1246 (2010).

[13] P. Chen, C.C. Chen, and Y.H. Peng, IEEE J. Solid-State Circuits, p.600(2010)

[14] S. Xie and W.T. Ng, "An All-digital Self-calibrated Temperature Sensor Implemented Using 65/60 nm FPGAs," to be submitted to IEEE Trans. Circuits and Systems II.