Self-Timed Thermally-Aware Circuits

David Fang, Filipp Akopyan, Rajit Manohar*
Computer Systems Laboratory
Electrical and Computer Engineering
Cornell University
Ithaca, NY 14853, U.S.A.

Abstract

Thermal management is becoming increasingly important in circuit designs with high power density. Circuits that overheat beyond specified operating conditions may suffer timing failures, or become damaged for various reasons, including thermal runaway. We present a novel application of a thermally sensitive circuit to automatically regulate the performance and power consumption of asynchronous circuits, with minimal implementation overhead, and free of interruption of operation.

1 Introduction

As the power density of modern integrated circuits continues to increase, power and temperature management become increasingly important and challenging [3]. In this paper, we demonstrate a mechanism that regulates the performance of asynchronous circuits using a simple thermally-sensitive circuit. Our approach does not require temperature measurement, rather, we leverage the temperature response of subthreshold devices to construct a temperature-sensitive delay element to directly regulate the speed of the system. Asynchronous circuits are capable of operating correctly in the presence of continuous and dynamic changes in delays [2], and can self-modulate their performance without interruption of operation.

2 Circuits

As a circuit heats up, the gate delays increase and the frequency naturally drops, reducing the circuit's dynamic power consumption and self-heat generation. However, the natural negative-feedback retardation of this self-heating rate is too weak to halt the increase in temperature [5].

In the subthreshold region, the source-drain current is exponentially dependent on temperature:
$I_D = I_0 \cdot \exp\left(\frac{V_{gs} \cdot q}{\zeta \cdot k \cdot T}\right)$ where I_D is the drain-source current, I_0 depends on channel width, channel length, diffusion constant of carriers, carrier density and electron charge [6], ζ is a nonideality factor, and T is temperature in K. We use the high temperature-sensitivity of the subthreshold transistors to construct a temperature-sensitive voltage source, shown boxed in Figure 1. Transistors M1 and M2 are biased differently to have contrasting thermal sensitivities—M1 operates in deep subthreshold (more temperature-sensitive), while M2 operates near-threshold—forming a temperature-sensitive, resistive voltage divider. The bias voltages and the sizes for M1 and M2 are tuned to achieve a desired temperature response for a given technology. Figure 1 shows a thermally-sensitive voltage source controlling the gate of a foot transistor in typical logic circuitry. As temperature increases, the foot transistor starves the current until the point where switching ceases, and the conditional staticizer retains the value of V_{int}.

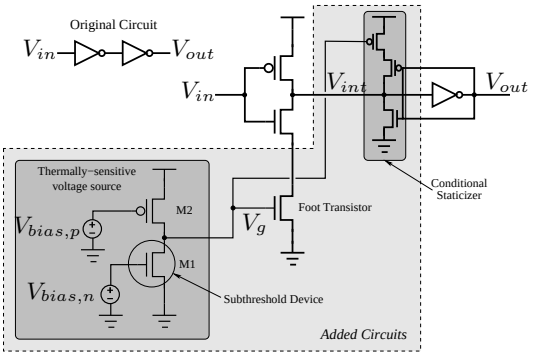

Figure 1. Temp.-sensitive delay element

We simulated the above temperature-dependent circuit using TSMC 0.18μm technology parameters, with 1.8 V nominal supply V_{dd}. Figure 2 shows the V_g switching sharply near 100°C. Above 113°C, V_{out} never switches—the current through the pull-down logic cannot overpower the pull-up staticizer, so the delay is infinite. Once the temperature drops sufficiently, operation will resume.

The peak current of the subthreshold transistor is on the order of tens of nA, comparable to leakage current for this

*E-mail: {fang,filipp,rajit}@csl.cornell.edu

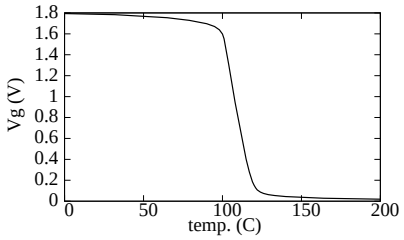

Figure 2. Foot Transistor V_g-T Dependence

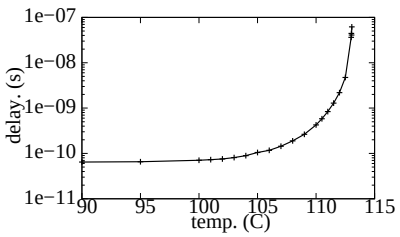

Figure 3. Delay through modified inverter

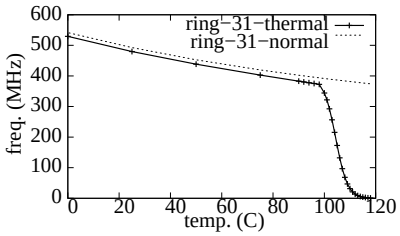

Figure 4. Freq. of ring oscillator vs. circuit temp.

technology, and is thus, negligible. For above-threshold devices, we use a conventional delay-derating model as a function of voltage and temperature [1].

3 Evaluation

In each scenario, we take the original digital circuit and then replace a few selected inverters with temperature-sensitive delay elements. The result is a circuit that regulates its speed based on the local temperature with minimal hardware overhead. Note that this applies to circuits of arbitrary complexity, because having a slow gate on a handshake cycle is sufficient to limit the local throughput of a complex asynchronous circuit. To demonstrate the effectiveness of our application of thermally-sensitive circuits we present examples of varying complexity.

Ring oscillator. A 31-stage ring oscillator with one stage modified to a thermally-sensitive inverter was simulated. The frequencies are plotted against temperature in Figure 4. As expected, the frequency quickly degrades beyond 100°C, where the delay grows exponentially. Before 95°C, the difference between the modified and unmodified oscillator shows that the scheme introduces minimal overhead.

FPGA. We simulated a 5x5 asynchronous FPGA (the design is described in [4]) running a function-block-intensive benchmark to demonstrate a circuit that reaches a self-heating equilibrium. We report the normalized throughput against the peak surface temperature in Table 1. The thermally-aware FPGA's average surface temperature stabilizes at around 100°C after 1 ms of simulated time at an operating frequency of 27% of the room-temperature throughput. In the same scenario, the FPGA without our thermal-aware modifications heats itself to destructive temperatures.

Automatic Dynamic Scheduling. Our technique can be used to dynamically schedule activity away from hot-spots. We designed a dynamic instruction scheduler that uses workload-driven non-deterministic scheduling to issue tasks to execution units A and B. Normally the scheduler issues an equal number of instructions to A and B. The two execution units were modified to use the thermally sensitive delay. We examined the impact of unit A operating at a higher temperature than B. Table 2 shows that beyond 100°C, unit A practically stops operating, yielding almost all of the computation work to unit B. Once unit A cools down, it will begin to request data from the dispatcher more frequently. Note that in this example, the dispatch circuit is unmodified.

Table 1. Normalized throughput (f) of thermal-aware FPGA at different temperatures

°C	25	45	80	89	94	97	100
f	1.00	0.85	0.70	0.64	0.60	0.49	0.27

Table 2. Processed data items in a fixed time window for dynamic scheduling.

Temp.-A (°C)	40	60	100	101
Instructions processed by A	46	43	12	8
Temp.-B (°C)	40	42	45	45
Instructions processed by B	46	44	49	48

4 Conclusion

We presented a simple thermally-sensitive circuit that can be used to regulate gate delays in digital circuits. We demonstrated a few applications of the technique in self-regulating asynchronous circuits, as well as in automatically dispatching instructions away from hot-spots without modifying the dispatch circuit. A small number of strategically placed thermally-sensitive gates is sufficient to regulate the local throughput in an asynchronous system with minimal circuit overhead and redesign effort in a way that prevents the circuit from exceeding its thermal budget.

References

[1] J. M. Daga, E. Ottaviano, and D. Auvergne. Temperature effect on delay for low voltage applications. In *Proc. DATE*, pages 680–685, 1998.

[2] Alain J. Martin. The limitations to delay-insensitivity in asynchronous circuits. In William J. Dally, editor, *Proc. ARVLSI*, pages 263–278. Massachusetts Institute of Technology, 1990.

[3] Dennis Sylvester and Himanshu Kaul. Future performance challenges in nanometer design. In *Proc. DAC*, pages 3–8, New York, NY, USA, 2001. ACM Press.

[4] J. Teifel and R. Manohar. Highly pipelined asynchronous FPGAs. In *Proc. FPGA*, February 2004.

[5] J. A. Tierno. *An Energy-Complexity model for VLSI computations*. PhD thesis, California Institute of Technology, 1995.

[6] Neil Weste and David Harris. *CMOS VLSI Design: A Circuits and Systems Perspective*. Addison-Wesley, 2005.