T_I-States: Power Management in Active Timing Margin Processors

Temperature inversion is a transistor-level effect that improves performance when temperature increases. This article presents a comprehensive measurement-based analysis of its implications for architecture design and power management using the AMD A10-8700P processor. The authors propose temperature-inversion states (T_I-states) to harness the opportunities promised by temperature inversion. They expect T_I-states to be able to improve the power efficiency of many processors manufactured in future CMOS technologies.

Yazhou Zu
University of Texas at Austin

Wei Huang
Indrani Paul
Advanced Micro Devices

Vijay Janapa Reddi
University of Texas at Austin

Temperature inversion refers to the phenomenon that transistors switch faster at a higher temperature when operating under certain regions. To harness temperature inversion's performance benefits, we introduce T_i-states, or *temperature-inversion states,* for active timing-margin management in emerging processors. T_i-states are frequency, temperature, and voltage triples that enable processor timing-margin adjustments through runtime supply voltage changes. Similar to P-states' frequency-voltage table lookup mechanism, T_i-states operate by indexing into a temperature-voltage table that resembles a series of power states determined by transistors' temperature-inversion effect. T_i-states push greater efficiency out of the underlying processor, specifically in active timing-margin-based processors.

T_i-states are the desired evolution of classical power-management mechanisms, such as P-states and C-states. This evolution is enabled by the growing manifestation of the transistor's temperature-inversion effect as device feature size scales down.

When temperature increases, transistor performance is affected by two factors: a decrease in both carrier mobility and threshold voltage. Reduced carrier mobility causes devices to slow down, whereas reduced threshold voltage causes devices to speed up. When supply voltage is low enough, transistor speed is sensitive to minute threshold voltage changes, which makes the second factor (threshold voltage reduction) dominate. In this situation, temperature inversion occurs.[1]

In the past, designers have safely discounted temperature inversion because it does not occur under a processor's normal operation. However, as transistor feature size scales down, today's processors are operating close to the temperature inversion's voltage

Published by the IEEE Computer Society

region. Therefore, the speedup benefit of temperature inversion deserves more attention from architects and system operators.

Figure 1a provides a device simulation analysis based on predictive technology models.[2,3] We use inflection voltage to denote the crossover point for temperature inversion to occur. Below the inflection voltage is the temperature-inversion region, in which circuits speed up at high temperature. Above the inflection voltage is the noninversion region, in which circuits slow down at high temperature. From 90 nm to 22 nm, the inflection voltage keeps increasing and approaches the processor's nominal voltage. This means temperature inversion is becoming more likely to occur in recent smaller technologies.

Our silicon measurement corroborates and strengthens this projection. The measured 28-nm AMD A10-8700P processor's inflection voltage falls within the range of the processor's different P-states. Figure 1b further illustrates temperature inversion by contrasting circuit performance in the inversion and noninversion regions. At 1.1 V, the circuit is slightly slower at a higher temperature while safely meeting the specified frequency, as expected from conventional wisdom. At 0.7 V, however, this circuit becomes faster by more than 15 percent at 80°C as a result of temperature inversion.

T_i-states harness temperature inversion's speedup effect by actively undervolting to save power. T_i-states exploit the fact that the faster circuits offered by temperature inversion add extra margin to the processor's clock cycle. It then calculates the precise amount of voltage that can be safely reduced to reclaim the extra margin. The undervolting decision for each temperature is stored in a table for runtime lookup.

T_i-states are instrumental because they can apply to almost all processors manufactured with today's technologies that manifest a strong temperature-inversion effect, including bulk CMOS, fin field-effect transistor (FinFET), and fully depleted silicon on insulator (FD-SOI). The comprehensive characterization we present in this article is based on rigorous hardware measurement, and it can spawn future work that exploits the temperature-inversion effect.

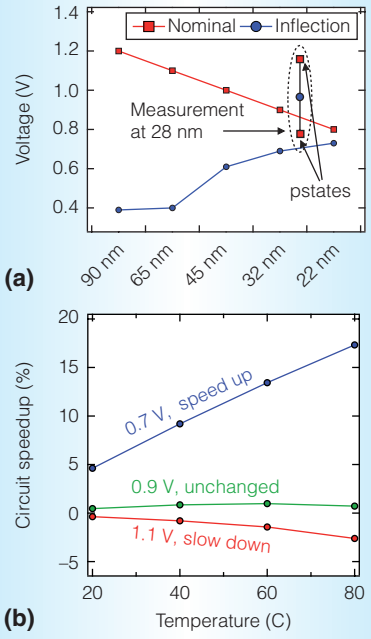

Figure 1. Temperature inversion is having more impact on processor performance as technology scales. (a) Temperature inversion was projected to be more common in smaller technologies as its inflection voltage keeps increasing and approaches nominal supply. (b) High temperature increases performance under low voltage due to temperature inversion, compared to conventional wisdom under high voltage.

Measuring Temperature Inversion

We measure temperature inversion on a 28-nm AMD A10-8700P accelerated processing unit (APU).[4] The APU integrates two CPU core pairs, eight GPU cores, and other system components. We conduct our study on both the CPU and GPU and present measurements at the GPU's lowest P-state of 0.7 V and 300 MHz, because it has strong temperature inversion. The temperature-inversion effect we study depends on supply voltage but not on the architecture. Thus, we expect the analysis on the AMD-integrated GPU to naturally extend to the CPU and other architectures as well for all processor vendors.

We leverage the APU's power supply monitors (PSMs) to accurately measure circuit speed changes under different conditions.[5]

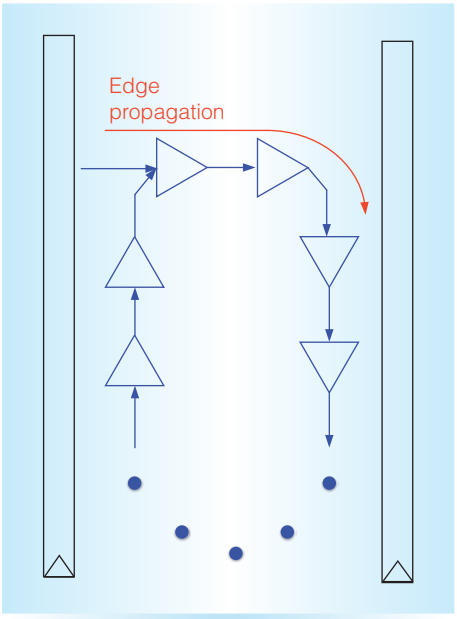

Figure 2. A power supply monitor (PSM) is a ring of inverters inserted between two pipeline latches. It counts the number of inverters an "edge" travels through in one cycle to measure circuit speed.

Figure 2 illustrates a PSM's structure. A PSM is a time-to-digital converter that reflects circuit time delay in numeric form. Its core component is a ring oscillator that counts the number of inverters an "edge" has traveled through in each clock cycle. When the circuit is faster, an edge can pass more inverters, and a PSM will produce a higher count output. We use a PSM as a means to characterize circuit performance under temperature variation. We normalize the PSM reading to a reference value measured under 0.7 V, 300 MHz, 0°C, and idle chip condition.

To characterize the effect of temperature inversion on performance and power under different operating conditions, we carefully regulate the processor's on-die temperature using a temperature feedback control system (see Figure 3). The feedback control checks die temperature measured via an on-chip thermal diode and adjusts the thermal head temperature every 10 ms to set the chip temperature to a user-specified value. Physically, the thermal head's temperature is controlled via a water pipe and a heater to control its surface temperature.

The Temperature-Inversion Effect

Temperature inversion primarily affects circuit performance. We first explain temperature inversion's performance impact with respect to supply voltage and temperature. We then extrapolate the power optimization potential offered by temperature inversion. Through our measurement, we make two observations: temperature inversion's speedup effects become stronger with lower voltage, and the speedup can be turned into more than 5 percent undervolting benefits.

Inversion versus Noninversion

We contrast temperature inversion and noninversion effects by sweeping across a wide operating voltage range. Figure 4 shows the circuit speed change under different supply voltages and die temperatures. Speed is reflected by the PSM's normalized output—a higher value implies a faster circuit. We keep the chip idle to avoid any workload disturbance, such as the di/dt effect.

Figure 4 illustrates the insight that the temperature's impact on circuit performance depends on the supply voltage. In the high supply-voltage region around 1.1 V, the PSM's reading becomes progressively smaller as the temperature rises from 0°C to 100°C. The circuit operates slower at a higher temperature, which aligns with conventional belief. The reason for this circuit performance degradation is that the transistor's carrier mobility decreases at a higher temperature, leading to smaller switch-on current (I_{on}) and longer switch time.

Under a lower supply voltage, the PSM's reading increases with higher temperature, which means the circuit switches faster (that is, the temperature-inversion phenomenon). The reason is because the transistor's threshold voltage (V_{th}) decreases linearly as temperature increases. For the same supply voltage, a lower V_{th} provides more drive current (I_{on}), which makes the circuit switch faster. The speedup effect is more dominant when supply voltage is low, because then the supply voltage is closer to V_{th}, and any minute change of V_{th} can affect transistor performance.

An "inflection voltage" exists that balances high temperature's speedup and slowdown effects. On the processor we tested,

the inflection voltage is between 0.9 V and 1 V. Around this point, temperature does not have a notable impact on circuit performance. Technology evolution has made more chip operating states fall below the inflection voltage (that is, in the temperature-inversion region). For the APU we tested, half of the GPU's P-states, ranging from 0.75 to 1.1 V, operate in the temperature-inversion region. Therefore, we must carefully inspect temperature inversion and take advantage of its speedup effect.

Active Timing Margin's Undervolting Potential

We propose to harness temperature inversion's speedup effect by reclaiming the extra pipeline timing margin provided by the faster circuitry. Specifically, we propose to actively undervolt to shrink the extra timing margin, an approach similar in spirit to prior active-timing-margin management schemes.[6] To explore the optimization space, we first estimate the undervolting potential using PSM measurement.

Figure 5 illustrates the estimation process. The *x*-axis zooms into the low-voltage region between 0.6 and 0.86 V in Figure 4 to give a closer look at the margin-reduction opportunities.

Temperature inversion's performance benefit becomes stronger at lower voltages, as reflected by the widening gap between 100°C and 0°C. At 0.7 V, the PSM difference between 100°C and 0°C represents the extra timing margin in the units of inverter delays. In other words, it reflects how much faster the circuits run at a higher temperature by counting how many more inverters the faster circuit can switch successively in one cycle. To bring the faster circuit back to its original speed, supply voltage needs to be reduced such that under a higher temperature the PSM can read the same value. We estimate the voltage reduction potential with linear extrapolation. At 0.7 V, the extra margin translates to a 46-mV voltage reduction, equivalent to 5 percent undervolting potential. See our original paper for more complete extrapolation results.[7]

Temperature-Inversion States

Based on our temperature-inversion characterization, we propose T_i-states to construct a safe undervolting control loop to reclaim the extra timing margin provided by temperature

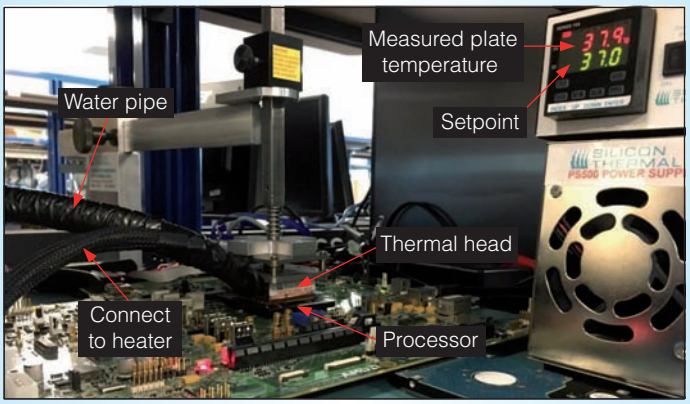

Figure 3. Temperature control setup. The thermal head's temperature is controlled via a water pipe and a heater. The water pipe is connected to an external chiller to offer low temperatures while the heater increases temperature to reach the desired temperature setting.

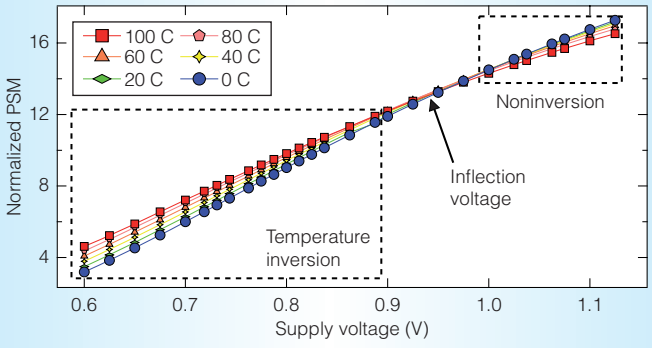

Figure 4. Circuit speed changes under different supply voltages and die temperatures. Temperature inversion happens below 0.9 V and is progressively stronger when voltage scales down.

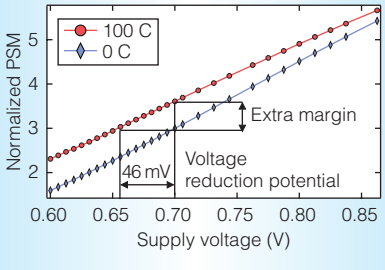

Figure 5. Temperature inversion happens below 0.9 V. It speeds up circuits, as reflected by larger PSM values under higher temperatures, and becomes stronger when voltage scales down.

inversion. In doing this, we must not introduce additional pipeline timing threats for reliability purposes, such as overly reducing timing margins or enlarging voltage droops caused by workload di/dt effects.

To guarantee timing safety, we use the timing margin measured at 0°C as the "golden" reference. We choose 0°C as the reference because it represents the worst-case operating condition under temperature inversion. Workloads that run safely at 0°C are guaranteed to pass under higher temperatures, because temperature inversion can make circuits run faster. Although 0°C rarely occurs in desktop, mobile, and datacenter applications, during the early design stage, timing margins should be set to tolerate these worst-case conditions. In industry, 0°C or below is used as a standard circuit design guideline.[8] In critical scenarios, an even more conservative reference of –25°C is adopted.

T_i-states' undervolting goal is to maintain the same timing margin as 0°C when a chip is operating at a higher temperature. In other words, the voltage T_i-state sets should always make the timing margins measured by the PSM match 0°C. Under this constraint, T_i-states undervolt to maximize power saving.

Algorithm 1 summarizes the methodology to construct T_i- states:

```
1:procedure GET REFERENCE  MARGIN
2:     set voltage and temperature to  reference
3:     for each training workload  do
4:           workloadMargin ← PSM measurement
5:           push RefMarginArr, workloadMargin
       returnRefMarginArr

6:procedure EXPLORE UNDERVOLT
7:     initVDD ← idle PSM extrapolation
8:     candidateVDDArr ← voltage around initVDD
9:     minErr ← MaxInt
10:    set exploration temperature
11:    for each VDD in candidateVDDArrdo
12:          set voltage to VDD
13:          for each training workload do
14:                workloadMargin ← PSM measurement
15:                push TrainMarginArr, workloadMargin
16:          err ← diff(RefMarginArr,TrainMarginArr)
17:          if err<minErr then
18:                minErr ← err
19:                exploreVDD ← VDD
       returnexploreVDD
```

The algorithm repeatedly stress tests the processor under different temperature-voltage environments with a set of workloads, and produces a temperature-voltage table that can be stored in system firmware.[9] At runtime, the system can index into this table to actively set the supply voltage according to runtime temperature measurement.

Algorithm 1 uses a set of workloads as the training sets to first get a tentative temperature-voltage mapping. We then validate this mapping with another set of test workloads.

During the training stage, Algorithm 1 first measures each workload's golden reference timing margin at 0°C using PSMs. The timing margin is recorded as the worst-case margin during the entire program run. Then, at each target temperature, Algorithm 1 selects four candidate voltages around the extrapolated voltage value as in Figure 5. The four candidate voltages are stepped through, and each workload's timing margin is recorded using PSMs. Finally, the timing margins at different candidate voltages are compared against the 0°C reference, and the voltage with the minimum PSM difference is taken as the target temperature's T_i-state voltage.

Table 1 shows the PSM difference compared with the 0°C reference across different candidate voltages for 20°C, 40°C, 60°C, and 80°C. The selected T_i-state voltages with the smallest difference are shown in bold in the table. For instance, at 80°C, 0.6625 V is the T_i-state, which provides around 5% voltage reduction benefits.

We observed from executing Algorithm 1 that a T_i-state's undervolting decision is independent of the workloads. It achieves the same margin reduction effects across all programs. This makes sense, because temperature inversion is a transistor-level effect and does not depend on other architecture or program behavior. This observation is good for T_i-states, because it justifies the applicability of the undervolting decision made from a small set of test programs to a wide range of future unknown workloads.

Figure 6 illustrates our observation. Going from 0°C to 80°C, temperature inversion offers more than 15 percent extra timing margin. T_i-states safely reclaim the extra margin by reducing voltage to 0.66 V. After voltage reduction, workload timing margins

Table 1. PSM error compared to the reference setting for different <temperature, voltage> configurations.

Candidate voltages (mV)	20°C	40°C	60°C	80°C	100°C
693.75	3.7%	—	—	—	—
687.50	**2.2**%	—	—	—	—
681.25	8.4%	**2.3**%	—	—	—
675.00	13.9%	5.3%	4.9%	—	—
668.75	—	9.5%	**2.5**%	—	—
662.50	—	13.5%	6.5%	**1.9**%	—
656.25	—	—	12.2%	5.6%	9.9%
650.00	—	—	—	9.3%	**5.1**%

*Bold type indicates the voltages with the smallest PSM difference.

closely track the baseline for all workloads. Overall, T_i-states can achieve 6 to 12 percent power savings on our measured chip across different temperatures.

Long-Term Impact

As CMOS technology scales to its end, it is important to extract as much efficiency improvement opportunity as possible from the underlying transistors. T_i-state achieves this goal with active timing-margin management. Exploiting slack in timing margins to improve processor efficiency will be ubiquitous, just as P-states and C-states have helped reduce redundant power in the past. We believe the simplicity of T_i-states and the insights behind them render a wide range of applicability. Our work brings temperature inversion's value from device level to architects and system managers, and opens doors for other ideas to improve processor efficiency.

Wide Range of Applicability

T_i-state is purely based on transistor's temperature-inversion effect and is independent of other factors. Temperature inversion is an opportunity offered by technology scaling, which makes it a free meal for computer architects. Therefore, T_i-state is applicable to chips made with today's advanced technologies, including bulk CMOS, FD-SOI, and FinFET (as we show in our original paper[7]). Many, if not all, processor architectures can benefit from it, whether they're CPUs, GPUs, FPGAs, or other accelerators.

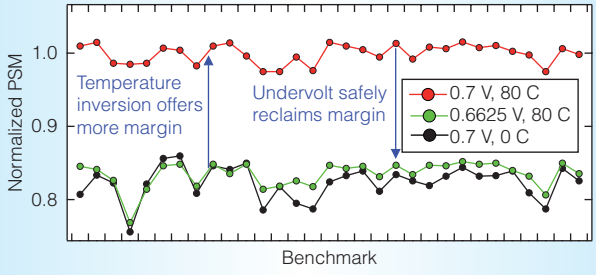

Figure 6. Temperature inversion's speedup effect offers extra timing margin at 80°C, as reflected by the elevated workload worst-case PSM. T_i-state precisely reduces voltage to have the same timing margin as under 0°C and nominal voltage, which achieves better efficiency and guarantees reliability.

T_i-state's design is succinct. Its main components are on-chip timing margin sensors, temperature sensors, and system firmware that stores T_i-state tables. A T_i-state's runtime overhead is a table lookup and a voltage regulator module's set command, which are minimal. Because chip temperature changes over the course of several seconds, a T_i-state's feedback loop has no strict latency requirement, which makes it easy to design, implement, and test.

Implications at Circuit, Architecture, and System Level

Our study conducted on an AMD A10-8700P processor focuses on a single chip made in planar CMOS technology. Going beyond current technology and across system scale, T_i-states will have a bigger impact in the future.

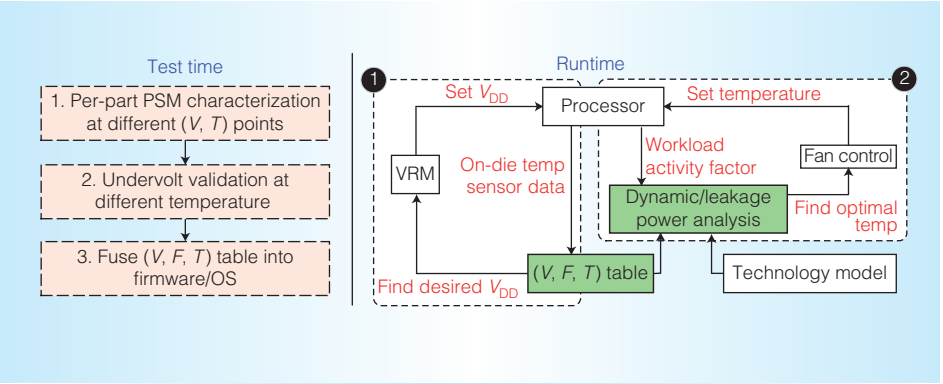

Figure 7. T_i-state temperature and voltage control: two loops work in synergy to minimize power. Loop 1 is a fast control loop that uses a T_i-state table to keep adjusting voltage in response to silicon temperature variation. Loop 2 is a slow control loop that sets the optimal temperature based on workload steady-state dynamic power profile.

Significance for FinFET and FD-SOI. FinFET and FD-SOI are projected to have stronger and more common temperature-inversion effects.[10,11] In these technologies, T_i-states have broader applicability and more benefits. Furthermore, the low-leakage characteristics of these technologies promise other opportunities for a tradeoff between temperature and power.

In our original paper, we provide a detailed FinFET and FD-SOI projection analysis based on measurements taken at 28-nm bulk CMOS. We find the 10-times leakage reduction capabilities make these technologies enjoy a higher operating temperature, because T_i-states reduce more V_{DD} under higher temperatures. The optimal temperature for power is usually between 40°C to 60°C, depending on workloads and device type. Thus, T_i-states not only reduce chip power itself for FinFET and FD-SOI but also relieve the burden of the cooling system.

System-level thermal management. Datacenters and supercomputers strive to make room temperature low at the cost of very high power consumption and cost. A tradeoff between cooling power and chip leakage power exists in this setting.[12] T_i-states add new perspective to this problem. First, we find that high temperature does not worsen timing margins, but actually preserves processor timing reliability because of temperature inversion. Second, T_i-states reduce power under higher temperatures, mitigating processor power cost. For FinFET and FD-SOI, the processor might prefer high temperatures around 60°C to save power, which further provides room for cooling power reduction.

Figure 7 shows a control mechanism that we conceived to synergistically reduce chip and cooling power. The test-time procedure and loop 1 is what the T_i-state achieves. In addition, loop 2 takes cooling system power into consideration and jointly optimizes fan and chip power together. Overall, temperature inversion and T_i-states enable an optimization space involving cooling power, chip power, and chip reliability.

Opportunity for near-threshold computing. Our measurement on a real chip shows that temperature inversion is stronger at lower voltages, reaching up to 10 percent V_{DD} reduction potential for a T_i-state at 0.6 V for our 28-nm chip. In near-threshold conditions as low as 0.4 V, temperature inversion will have a much stronger effect and will offer much larger benefits. In addition to power reduction, a T_i-state can be employed to boost the performance of near-threshold chips by overclocking directly to exploit extra margin. Extrapolation similar to Figure 4 shows overclocking potential is between 20 and 50 percent with the help of techniques that mitigate di/dt effects.[5]

Temperature inversion offers a new avenue for improving processor efficiency.

On the basis of detailed measurements, our article presents a comprehensive analysis of how temperature inversion can alter the way we do power management today. Through the introduction of T_i-states, we show that active timing margin management can be successfully applied to exploit temperature inversion. Applying such optimizations in the future will likely become even more important as technology scaling continues. We envision future work that draws on T_i-states to enhance computing systems across the stack and at a larger scale. MICRO

References

1. C. Park et al., "Reversal of Temperature Dependence of Integrated Circuits Operating at Very Low Voltages," *Proc. Int'l Electron Devices Meeting* (IEDM), 1995, pp. 71–74.
2. D. Wolpert and P. Ampadu, "Temperature Effects in Semiconductors," *Managing Temperature Effects in Nanoscale Adaptive Systems*, Springer, 2012, pp. 15–33.
3. W. Zhao and Y. Cao, "New Generation of Predictive Technology Model for Sub-45 nm Early Design Exploration," *IEEE Trans. Electron Devices*, vol. 53, no. 11, 2006, pp. 2816–2823.
4. B. Munger et al., "Carrizo: A High Performance, Energy Efficient 28 nm APU," *J. Solid-State Circuits* (JSSC), vol. 51, no. 1, 2016, pp. 1–12.
5. A. Grenat et al., "Adaptive Clocking System for Improved Power Efficiency in a 28nm x86–64 Microprocessor," *Proc. IEEE Int'l Solid-State Circuits Conf.* (ISSCC), 2014, pp. 106–107.
6. C.R. Lefurgy et al., "Active Management of Timing Guardband to Save Energy in POWER7," *Proc. 44th IEEE/ACM Int'l Symp. Microarchitecture* (MICRO), 2011, pp. 1–11.
7. Y. Zu, "T_i-states: Processor Power Management in the Temperature Inversion Region," *Proc. 49th Ann. IEEE/ACM Int'l Symp. Microarchitecture* (MICRO), 2016; doi:10.1109/MICRO.2016.7783758.
8. *Guaranteeing Silicon Performance with FPGA Timing Models*, white paper WP-01139-1.0, Altera, Aug. 2010.
9. S. Sundaram et al., "Adaptive Voltage Frequency Scaling Using Critical Path Accumulator Implemented in 28nm CPU," *Proc. 29th Int'l Conf. VLSI Design and 15th Int'l Conf. Embedded Systems* (VLSID), 2016, pp. 565–566.
10. W. Lee et al., "Dynamic Thermal Management for FinFET-Based Circuits Exploiting the Temperature Effect Inversion Phenomenon," *Proc. Int'l Symp. Low Power Electronics and Design* (ISLPED), 2014, pp. 105–110.
11. E. Cai and D. Marculescu, "TEI-Turbo: Temperature Effect Inversion-Aware Turbo Boost for FinFET-Based Multi-core Systems," *Proc. IEEE/ACM Int'l Conf. Computer-Aided Design* (ICCAD), 2015, pp. 500–507.
12. W. Huang et al., "TAPO: Thermal-Aware Power Optimization Techniques for Servers and Data Centers," *Proc. Int'l Green Computing Conf. and Workshops* (IGCC), 2011; doi:10.1109/IGCC.2011.6008610.

Yazhou Zu is a PhD student in the Department of Electrical and Computer Engineering at the University of Texas at Austin. His research interests include resilient and energy-efficient processor design and management. Zu received a BS in microelectronics from Shanghai Jiao Tong University of China. Contact him at yazhou.zu@utexas.edu.

Wei Huang is a staff researcher at Advanced Micro Devices Research, where he works on energy-efficient processors and systems. Huang received a PhD in electrical and computer engineering from the University of Virginia. He is a member of IEEE. Contact him at wein.huang@amd.com.

Indrani Paul is a principal member of the technical staff at Advanced Micro Devices, where she leads the Advanced Power Management group, which focuses on innovating future power and thermal management approaches, system-level power modeling, and APIs. Paul received a PhD in electrical and computer engineering from the Georgia Institute of Technology. Contact her at indrani.paul@amd.com.

Vijay Janapa Reddi is an assistant professor in the Department of Electrical and Computer Engineering at the University of Texas at Austin. His research interests span the definition of computer architecture, including software design and optimization, to enhance mobile quality of experience and improve the energy efficiency of high-performance computing systems. Janapa Reddi received a PhD in computer science from Harvard University. Contact him at vj@ece.utexas.edu.