TOWARDS TEMPERATURE-INSENSITIVE NANOSCALE CMOS CIRCUITS WITH ADAPTIVELY REGULATED VOLTAGE POWER SUPPLIES

Ming Zhu[1], Yingtao Jiang[1], Mei Yang[1], Xiaohang Wang[2]

[1]Department of Electrical and Computer Engineering
University of Nevada Las Vegas, Las Vegas, NVUSA 89154
[2]School of Software, South China University of Technology, Guangzhou, China

ABSTRACT

In this paper, we show that the temperature-induced performance drop seen in nanoscale CMOS circuitscan be tackled by powering the circuits with adaptively regulated voltage power supplies. Essentially, when temperature rises, the supply voltage will be bumped up to offset otherwise performance degradation. To avoid thermal over-drift as chip temperature exceeds its operation range, a voltage limiteris integrated into the proposed power supply to cap the supply voltage. Using this proposed adaptive voltage source to power individual CMOS logic gates and/or subsystems will free the chips from using expensive high-precision temperature sensors for thermal management and performance tuning. Experiments on various benchmark circuits, which are implemented with a 45nm CMOS technology, have confirmed that the circuit delay variation can be reduced to 15%~30% over a wide temperature range (0℃ to 90℃), a sharp contrast to the large delay variations(50%~75%)observed in most IC designs where a constant power supply is employed.

KEYWORDS

High performance VLSI circuits; temperature-insensitive; voltage control; power supply.

1. INTRODUCTION

Variations in power supply voltages (V_{pp}) and temperatures (T) have strong implications on the delay of a CMOS circuit, and even may cause the chip to fail. In general, as chip temperature rises, two effects are observed: (i) the descending mobility of the transistor carriers, which contributes to the increase of the circuit propagation delay, and (ii) the decreasing of the absolute value of the threshold voltage (V_{th})of a transistor, which leads to a better delay performance. Experiments show that when both thermal effects come into play, the delay of a circuit operating at 90°Ccould be twice of that at 0°C. Moreover, such sheer thermal-induced delay performance penalty can vary significantly from one chip area to another [1], due to the imbalanced utilization and diversity of circuitry at different sections. As a result, ensuring the performance resilience against a wide range of temperature variations has become one of the greatest challenges facing nano scale VLSI circuit designs[2],

One most effective way to combat this temperature-induced delays is by changing the V_{pp} of the circuits. Along this line, two different approach shave been considered in the literature. The first approach is based on the fact that if$V_{pp}$is pulled up, it can offset the circuit delay introduced by

DOI : 10.5121/vlsic.2017.8301

the rise of temperature. Such approach requires the use of a power manager, with a look-up table (LUT)[4], to alter the V_{pp}according to the readings from the embedded temperature sensors[3] in the chip. By doing so, delay variations of primary Boolean logic gates could be confined within 20% as the chip($0.25\mu m$ technologies) temperature varies from 0℃ to 90℃[5].Furthermore, the work in[6] harnessed PVT sensors in multiple voltage and frequency domains of the chip, and each region had its ownV_{pp}that could be adjusted individually. One big downside of all these designs [4]-[6] is that they require high precision temperature sensors and abundant power control circuits, which are not easy to come by.

Another approach to deal with the temperature-induced delay variations is through powering the circuit at the zero-temperature-coefficient (ZTC) voltage level [2][7]. Essentially, by doing so, the thermal effects on the carrier mobility and MOSFET threshold are approximately canceled out, and thus, performance variations of the logic circuits could be negligible over a wide temperature range, from 25℃ to 125℃in [2]. Unfortunately, this ZTC voltage is usually close to the threshold voltage of MOSFETs, which can be problematic to high speed circuit designs. Rather, this approach is more suitable for low power, but less performance-critical applications.

In what follows, we will examine the temperature-delay relationship of basic CMOS logic gates in Section 2, through theoretical analysis and circuit simulations. Such relationship suggests the use of an adaptive power supply as detailed in Section 3.In light of a complementary to absolute temperature (CTAT) current source shown in [8][9], the proposed adaptive power supply source, to a first order approximation, can be used to build temperature-insensitive logic circuits and subsystems. In addition, a voltage limiter is integrated into the proposed power source to limit the output voltage after the temperature rises above a certain point. With this adaptive power supply, logic circuits experience much smaller circuit delay variations over a wide temperature range (Section 4), but no high-precision temperature sensors are required. Finally, the conclusion is drawn in Section 5.

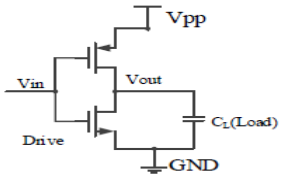
2. DELAY OF BASIC CMOS LOGIC GATES VS. TEMPERATURE AND VOLTAGE

Principally, all primitive CMOS logic gates can be conceptually collapsed to inverters with the same delay performance. Without loss of generality, the propagation delay model of a CMOS inverter (Figure 1),related to the time spent on charging or discharging its capacitive load [10], can be expressed as:

$$Delay = \frac{Q}{I} = \frac{C_L V_{pp}}{I} \qquad (1)$$

where $I = \left(\frac{W}{L}\right)\mu(T)C_{ox}\left[(V_{in} - V_{th})V_{out} - \frac{1}{2}V_{out}^2\right]$, for $V_{out} \le V_{in} - V_{th}$

$$I = \left(\frac{W}{L}\right)\mu(T)C_{ox}(V_{in} - V_{th})^2, \text{ for } V_{out} > V_{in} - V_{th}$$



1: Gate delay model of an inverter with a capacitive load.

where Q is the amount of charge to drive the capacitive load to logic 1/0 (i.e. V_{pp}/V_{GND}); I is the charging/discharging current through the transistor;C_L is the load capacitance; W and L are the channel width and length of the transistor, respectively; $\mu(T)$ is the carrier mobility, roughly proportional to T^{-2}[10]; $V_{th}(T)$ decreases with the growth of T in a nearly linear fashion [10][11]; and C_{ox} is the unit gate capacitance of the MOSFET.

If the input is considered as a step signal, a simplified approximation of Eq. (1) can be given as:

$$Delay = \frac{C_L}{(W/L)C_{ox}\lambda}\frac{V_{pp}}{\mu(T)\left(V_{pp}-V_{th}(T)\right)^2} \qquad (2)$$

Where λ is a coefficient introduced to simplify Eq.(1). Therefore, once the process and design parameters (e.g. W, L, C_{ox}) are fixed, we can only regulate V_{pp} as a function of T to offset the thermal effect on the propagation delay. Given that $\mu(T) \propto T^{-2}$, and $V_{th}(T)$ varies negligibly in a range of 0℃ to 90℃[10], we can treat $V_{th}(T)$ as a constant, V_{th}, and obtain the derivative of the delay described in Eq. (2) with respect to temperature, given as Eq.(3), where A is a positive, process-related, temperature-independent coefficient.

$$\frac{\partial}{\partial T}Delay = \frac{-A\left(\frac{\partial \mu(T)}{\partial T}\left(V_{pp}(T)-V_{th}\right)V_{pp}(T)+\mu(T)\frac{\partial V_{pp}(T)}{\partial T}\left(V_{pp}(T)+V_{th}\right)\right)}{\mu^2(T)\left(V_{pp}(T)-V_{th}\right)^3} \qquad (3)$$

Eq. (3) suggests that,in principle, proper adjustment of V_{pp} (e.g $V_{pp} \propto T^{\beta}, \beta > 0$)can make$\partial Delay/\partial T$ equal to 0, i.e., the temperature-induced delay can be eliminated. This effect is examined through simulations of a CMOS inverter implemented with a 45nm CMOS technology. Since the supply voltage of the 45nm technology is usually in range of 0.8V to 1.0V, we run simulations with a supply voltage range from 0.3V (slightly higher than V_{th}) to 1.2V. The delays versus temperatures at different V_{pp} levels are plotted in Figure 2, and the delay change rates with respect to temperatures at different V_{pp} levels are also reported in Table 1. The results reveal that when V_{pp} is set to be low (around 0.3V in Figure 2), the circuit delay at 0℃ is as high as 800 ps, and it drops as the temperature increases. As V_{pp} is set to be slightly higher, reaching the ZTC voltage [2][7], about 0.32V in our experiments, the circuit delay is approximately insensitive to the temperature changes. Once V_{pp} continues to rise to an even higher level (e.g. 1.0V, as is usually employed in high-speed circuits), the delay drops significantly (6.1ps at 0℃), but it rises rapidly as temperature increases (the delay actually is more than doubled at 90℃, compared to that at 0℃).

By studying the delays at different temperatures and V_{pp} levels in Figure 2, one can see that, if $V_{pp}$is set to be around 0.7V at 0℃, 0.8V at 30℃, 1.0V at 60℃, 1.2V at 90℃, respectively, (i.e., a net change of 0.5V for V_{pp} when the temperature rises from 0℃ to 90℃), the inverter's delay is almost independentof the temperature variations (about 10ps across the temperature range as shown in the red dashed line in Figure 2). Similarly, if we relax the performance requirement of the circuit, we can setV_{pp}properly so that the inverter delay is nearly unchanged from 0℃ to 90℃ at about 15ps (the green dotted line in Figure 2), while V_{pp} merely changes in a range from 0.6V to 0.8V. Agreeing to Eq. (3), these observations indicate that by using a temperature-adaptive

voltage power supply source, which is detailed in the next section, delay variations can be well controlled.

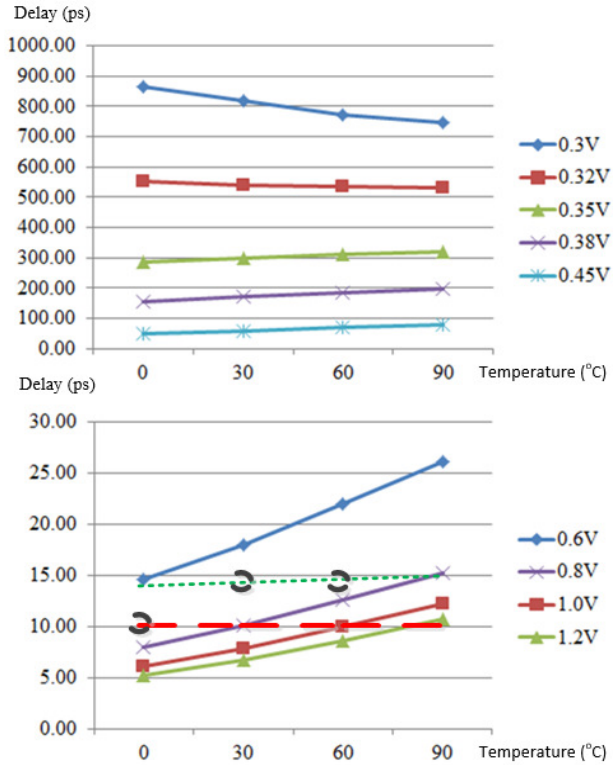


Figure 2 Inverter delay (ps) against temperature under different V_{pp} levels

Table 1 Change rates of delay (ps) vs. temperature under various V_{pp}

V_{pp} (V)	Avg. $\Delta t/\Delta T$ (ps/10°C)		Avg. $\Delta t/\Delta T$ var. (%/10°C).	
	At 0°C	At 90°C	At 0°C	At 90°C
0.30	-22.4	-7.4	-2.7	-0.9
0.32	-5.5	-1.5	-1.0	-0.3
0.35	3.2	4.0	1.1	1.3
0.38	4.3	3.7	2.5	2.0
0.45	2.9	3.3	4.8	4.9
0.60	1.1	1.4	5.8	6.4
0.80	0.6	0.9	6.0	7.1
1.00	0.4	0.8	5.6	7.7
1.20	0.4	0.7	5.8	8.4

3. SELF-ADAPTIVE, TEMPERATURE-AWARE VOLTAGE POWER SUPPLY CIRCUIT

To offset the performance penalty introduced by the rising temperature, as alluded in the previous section, V_{pp}needs to be slightly increased. In light of the linearity between the output current and temperature in the CTAT [8][9] current source, a self-adaptive, temperature-aware voltage power supply with a voltage limiter(Figure 3, where V_{DD} refers to the external DC voltage source, and

V_{pp} is the regulated temperature-adaptive voltage source that actually powers the digital logic circuits)is proposed. Compared to the designs in [4]-[6], no high-precision temperature sensors or redundant power supply managers are required. In addition, the voltage limiter restrains V_{pp}within a certain predefined range; that is, when the temperature exceeds certain level, V_{pp} will stop climbing, preventing the circuit from overheating due to the overrun of V_{pp}. In order to optimize the voltages at each node for a large dynamic swing over the entire temperature range of interest, we first study the equivalent impedance of a PMOS/NMOS, with its drain connecting to its gate (i.e.$V_{gs} = V_{ds}$), as given by Eq.(4).

$$Z_{eq} = \frac{V_{ds}}{I_{ds}} = \frac{A}{\mu(T)\left(V_{gs} - 2V_{th} + V_{th}^2/V_{gs}\right)} \qquad (4)$$

Where I_{ds} is the current from the drain to the source through the MOS channel and *A* is a constant coefficient that only relates to the manufacture process. Eq. (4) indicates that the equivalent impedance of the PMOS/NMOS increasesroughly proportional to T^2.Based on this observation, resistors, which have a relatively low (linear) temperature sensitivity [10], can be employed to replace some PMOS transistors where$V_{gs} = V_{ds}$. That is, impedances Z1, Z4, Z5 and Z6 in Figure 3 could be implemented as either resistors or PMOS/NMOS (Figure 3). By doing so, we can expect $V_{pp}$to be truly adaptive to the temperature change. It shall be noticed that the resistance values and the size of each transistor need to be carefully chosen to match the impedance of the logic circuits or subsystems powered by V_{pp} , so that the V_{pp} value at a high temperature, say 90℃, could drive the logic circuit with approximately the same delay as that of V_{pp} value at a low temperature, say 0℃. In addition, the output impedance induced by Z7 and MN6 should be small (i.e. relatively large width of MN6), so that this CTAT-like voltage power supply could drive a complicated logic circuit and/or a subsystem.

On the other hand, since the output voltage of this CTAT-like power supply is designed to rise as the temperature climbs, this increasing voltage output may exacerbate the circuit power consumption. To prevent this problem from happening, an impedance (Z8) and a diode (D1)are added to function as a voltage limiter(Figure 3). This voltage limiter has little effect when V_{pp} is low, but once V_{pp} exceeds a certain voltage that turns D1 on, V_{pp}will be capped. This voltage limiter helps protect the circuit from overheating, at a cost of performance degradation only at extremely high temperature (e.g. 120℃or higher).

In this paper, we provide an area-efficient (not necessarily delay optimized) CTAT-like voltage supply (Figure 4). In this case, Z1, Z4, Z5 and Z6 in Figure 3 are still implemented using PMOS transistors, yet all the widths of the PMOS and NMOS devices are set fixed, except for the width of MN6 and the resistance of R7(i.e. Z7 in Figure 3) that require manual adjustment during circuit design/layout process to match the impedances of different logic circuits powered by V_{pp}. Meanwhile, two NMOS are cascaded as the voltage limiter by connecting the body of NMOS to V_{pp} and its drain (together with the source and the gate) to the ground. As a result, each NMOS device is forward-biased, from the P-type body to the N-type drain/source, working as a diode to set a limit on the output voltage V_{pp}.

In the circuit shown in Figure 4, only MP3 operates in the saturation mode, while MP7,MP8, MN2 and MN4 are all in the linear mode. Since the impedance of PMOS changes more rapidly

than that of NMOS[10], voltages at nodes 2, 5, 6 and 7 will decrease with the increase of temperature, and thus, the current through MN6 will actually decrease, leading to an increased output voltage, V_{pp}. In addition, transistor pairs (MN1, MN2), (MN3, MN4) and (MN5, MN6) can help magnify the output voltage changes with respect to the temperature variations.

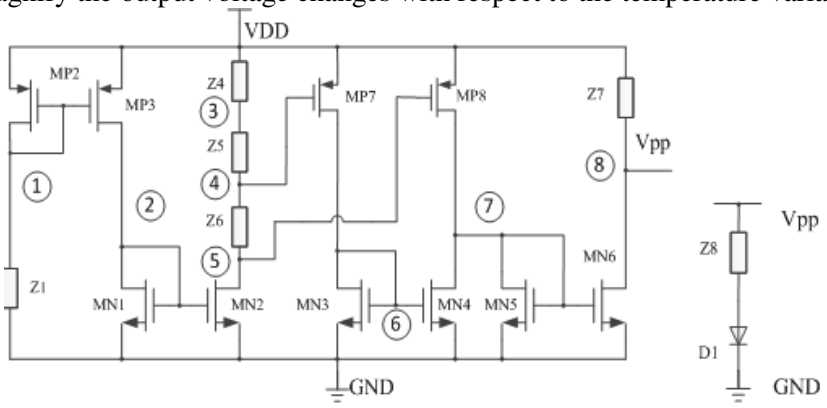


Figure 3 The proposed CTAT-like voltage power supply (left) and thevoltage limiter (right)

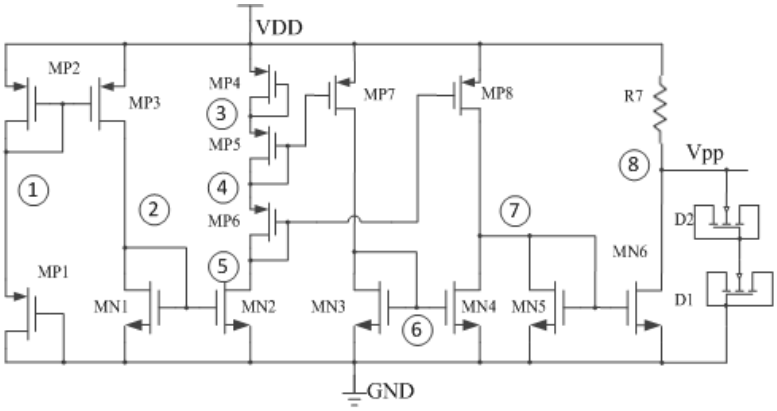


Figure 4 A general-use CTAT-like voltage power supply with voltage limiter

In a simple term, with the proposed CTAT-like power supply, higher V_{pp} can be achieved as temperature increases, so that the delay can be held to a relatively constant level. This voltage saturates as temperature reaches a preset level determined by the voltage limiter.

4. SIMULATION RESULTS

To verify the performance advancement of the proposed technique in terms of delay variations over a wide temperature range, we run HSPICE simulations on a number of benchmark circuits that are implemented using a 45nm CMOS technology. Each benchmark circuit is powered by one of the three different supplies (Figure 5): 1) a 1.0V constant power, 2) the proposed CTAT-like temperature-adaptive power supply without voltage limiter (Figure 3, denoted as CTAT-like thereafter), and 3) the proposed CTAT-like temperature-adaptive power supply with voltage limiter (Figure 4, denoted as CTAT_vr). We then measure the propagation delays and variations of these benchmarks operating in all scenarios, respectively, with a temperature swing from 0℃ to 180℃.

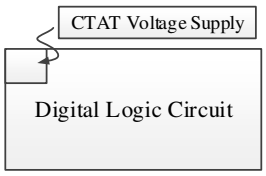


Figure 5 The circuit architecture that includes a power supply and a logic circuit.

Table 2 Widths (nm) of PMOS/NMOS in Figure 4

MOS	W	MOS	W	MOS	W	MOS	W	MOS	W
MP1	180	MP4	180	MP7	720	MN2	360	MN5	90
MP2	180	MP5	180	MP8	180	MN3	90		
MP3	180	MP6	180	MN1	90	MN4	90		

We size the CTAT-like power supply circuits in the way that the delays of a benchmark circuit with the proposed CTAT-like power supply (0℃ to90℃)match the delay of the same circuit when powered with a constant 1.0V at 30℃ (the baseline implementation).

We setV_{DD} in Figure 3 and Figure 4 to be 2V, and the widths of the PMOS/NMOS transistors adopted in this power supply are listed in Table 2. The width of MN6 can beset somewhere between 360nm and 2160nm, while the resistance of R7 is in the range of 500Ω or 2kΩdepending on the type of logic circuit thatV_{pp} is applied to (e.g. R7 can be about 600 Ω and MN6 at 2160nm for all the benchmark circuits used in the following simulations, including inverters, FAs, DFFs and ISCAS-85 benchmarks). By doing so, V_{pp}in Figure 3 could fall into the range of about 0.9V to 1.5V, while V_{pp} in Figure 4 are limited up to 1.4V, so that the logic circuit it powers could maintain a relatively uniform delay performance across the temperature range of0℃to 90℃. With a total of 14 PMOS/NMOS transistors and a resistor, this circuit occupies an area of about 0.2mm^2 in a 45nm technology.

We first run simulations and measure the propagation delays on simple logic circuits, such as a single-stage inverter, a three-stage cascaded inverter, one-bit full adder (FA) and a 4-bit ripple carry adder (RCA), powered by the three voltage supplies, respectively. The simulation results are shown in Figure 6 and Figure 7. One can see that, when these simple circuits are powered by a constant 1.0V power supply, delay variations exceed 60% when temperature goes from 0℃ to 90℃ (Figure 6), while if they are powered by the proposed CTAT-like voltage supply, delay variations are down to be about 15%~30%. On average, the proposed CTAT-like power supply can suppress the temperature-induced delay variation by 40%and more. Note that the proposed power supplies with and without a voltage limiter deliver almost indistinguishable performance for temperature up to90℃. However, the circuit performances with the voltage limiter degrade quickly once exceeding120℃, due to the blocking on V_{pp}'s growth, yet still outperform the ones powered by constant 1V.

We also run experiments over more complex circuits and subsystems, including a d-type flip-flop (D-FF) and three ISCAS-85 benchmarks (Figure 7), including C6288, which is a 16-by-16 binary multiplier, C499, a 32-bit Single-Error-Correcting (SEC) circuit, and C432, a 27-channel interrupt controller; all of these benchmark circuits are also powered by the three voltage supply

scenarios, respectively. The results shown in Figure 6 and Figure 7 demonstrate that, of all the four circuits, if they are powered by the proposed CTAT-like voltage power, the thermal-related delay variations range from about 15% to30%, whereas the delays of circuits powered by the uniform power supply of 1.0V vary between 50% and 70%.Meanwhile, the voltage limiter also capsV_{pp} when temperature exceeds 120℃, but barely affects the performance at 0℃ to 90℃. It should be noted that, CTAT_vr may provide a lower circuit delay variation than CTAT-like because the absolute circuit delay at 30℃ powered by CTAT_vr is slightly higher than that supplied by CTAT-like.

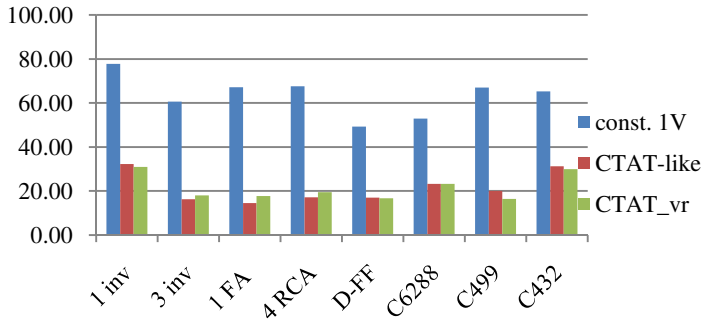


Figure 6 Delay variations (%) of inverters, adders, D-FF and ISCAS-85 benchmark circuits over a temperature range of0℃ to 90℃.

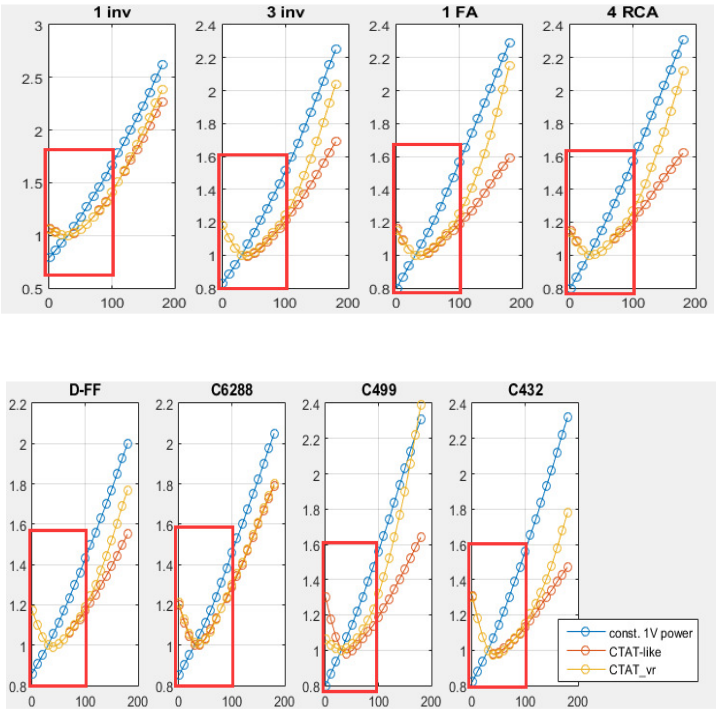


Figure 7 Delays (normalized over the baseline) of a single inverter, an inverter chain of three cascaded inverters, a 1-bit FA, and a 4-bit RCA, D-FF and ISCAS-85 benchmark circuits.

While using the proposed adaptive power supply circuit can help manage the temperature-induced delays in CMOS circuits, it comes with a cost. The proposed CTAT-like voltage power has a static leakage current, ranging from 1mA to 2mA, depending on the temperature and digital circuits it powers. In the worst scenario, if $V_{DD} = 2\text{V}$ as adopted in the simulations, the static power could be as high as 4mW. Nevertheless, as the temperature rises, the impedances of CTAT-like power supply tends to increase as well, leading to lower static current and consequently, lower static power consumption (static current drops from 1.5mA at 0℃ to 1.0mA at 90℃), which actually helps prevent the operation temperature from continuing to rise. In addition, if power gating techniques are applied, such static power consumption can be minuscule.

5. CONCLUSIONS

Due to the continuous scaling of integrated circuits to deep nanoscale, temperature variations could have substantial impact on the delay of a logic circuit. By exploring the temperature-voltage-delay relationship, in this paper, we have demonstrated that delay variations resulting from temperature changes can be significantly reduced using the proposed self-adaptive power supply. Of the benchmark circuits adopted in the experiments, if they were driven by a constant voltage power supply, the delay variations could be as high as 50 to 75% over a 90℃ temperature range; however, for the same circuits but powered by the proposed CTAT-like voltage power supply, the delay variations dropped significantly, to a significantly lower level, between15% and 30%.

ACKNOWLEDGEMENTS

This material is based, in part, upon work supported by the Science and Technology Research Grant of Guangdong Province No. 2016A010101011.

REFERENCES

[1] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, “Parameter Variations and Impact on Circuits and Microarchitecture,” Design Automation Conf., Anaheim, CA, Jun. 2003, pp. 338–342.

[2] R. Kumar and V. Kursun, “Temperature Variation Insensitive Energy Efficient CMOS Circuits in a 65nm CMOS Technology,” IEEE Int'l. Midwest Symp. Circuits and Systems, San Juan, PR, Aug. 2006, pp. 226–230.

[3] S. W. Chen, M. H. Chang, W. C. Hsieh, and W. Hwang, “Fully On-Chip Temperature, Process, and Voltage Sensors,” IEEE Int'l. Symp. Circuits and Systems, Paris, France, May 2010, pp. 897–900.

[4] M. Elgebaly, A. Fahim, I. Kang, and M. Sachdev, “Robust and Efficient Dynamic Voltage Scaling Architecture,” IEEE Int'l. SoC Conf., Portland, OR, Sept. 2003, pp. 155–158.

[5] A. Bellaouar, A. Fridi, M. J. Elmasry, and K. Itoh, “Supply Voltage Scaling for Temperature Insensitive CMOS Circuit Operation,” IEEE Trans. Circuits and Systems. II Analog and Digital Signal Process., vol. 45, no. 3, pp. 415–417, Mar. 1998.

[6] M. H. Chang et al., “Near-/Sub-Vth Process, Voltage, and Temperature (PVT) Sensors with Dynamic Voltage Selection,” IEEE Int'l. Symp. Circuits and Systems, Beijing, China, May 2013, pp. 133–136.

[7] F. Shoucair, "Design Consideration in High Temperature Analog CMOS Integrated Circuits," IEEE Trans. Compon. Hybrids Manuf. Technol., vol. 9, no. 3, pp. 242–251, Sept. 1986.

[8] S. V. Shinde, "PVT Insensitive Reference Current Generation," Int'l. MultiConference of Engineers and Computer Scientists, Hong Kong, China, Mar. 2014, pp. 690–694.

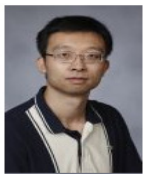
[9] S. Tang, S. Narendra, and V. De, "Temperature and Process Invariant MOS-based Reference Current Generation Circuits for Sub-1V Operation," Int'l. Symp. Low Power Electronics and Design, Seoul, South Korea, Aug. 2003, pp. 199–204.

[10] D. Wolpert and P. Ampadu, Managing Temperature Effects in Nanoscale Adaptive Systems. Springer-Verlag New York, 2012.

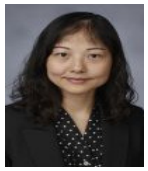
[11] N. Goel and A. Tripathi, "Temperature effects on Threshold Voltage and Mobility for Partially Depleted SOI MOSFET," Int'l. J. Comput. Appl., vol. 42, no. 21, pp. 56–58, Mar. 2012.

AUTHORS

Ming Zhu received his M.S. in electrical engineering from the University of Nevada, Las Vegas, in 2013.He is currently pursuing his Ph.D. degree at the same university. His research interests include system-on-chip design, statistic static timing analysis and error resilient computer systems.



Yingtao Jiang received his Ph.D. in computer science from the University of Texas at Dallas in 2001. He is currently a Professor in the Department of Electrical and Computer Engineering at the University of Nevada Las Vegas. His research interests include algorithms, computer architectures, biomedical signal processing and instrumentations, wireless communications and security, nuclear sensors and real-time, portable analytical instrument, and renewable energy.



Mei Yang received her Ph.D. degree in computer science from the University of Texas at Dallas, TX, in 2003. She is currently an Associate Professor in the Department of Electrical and Computer Engineering at University of Nevada, Las Vegas, NV. Her research interests include computer architectures, on-chip interconnection networks, embedded systems, and networking.

Xiaohang Wang received the B.Eng. and Ph.D. degree in communication and electronic engineering from Zhejiang University, in 2006 and 2011. He is currently an associate professor at South China University of Technology. He was the receipt of PDP 2015 and VLSI-SoC 2014 Best Paper Awards. His research interests include many-core architecture, power efficient architectures, optimal control, and NoC-based systems.