

A Dynamic Timing Control Technique Utilizing Time Borrowing and Clock Stretching

Kwanyeob Chae, Saibal Mukhopadhyay, Chang-Ho Lee, and Joy Laskar

Georgia Institute of Technology, Atlanta, GA 30332

Abstract- In this paper, a dynamic timing control technique employing a time-borrowing flip-flop with a time-borrowing detection and a clock shifter is presented to prevent timing errors of a system with a minimized performance penalty. The proposed flip-flop allows time borrowing during a time-borrowing window (T_{BW}) on critical paths and generates a time-borrowing detection signal used by the clock shifter to stretch the clock period by T_{BW} . This makes the system delay-error tolerant at a lower voltage or a higher frequency without any error management. To validate the proposed technique, we designed a prototype in a 180-nm CMOS technology. At a 10% activation probability of critical paths, the measurement results show a power reduction of up to 22% (at the same clock frequency) or an operating frequency increase of up to 10% (at the same power) compared to those of a conventional design.

I. INTRODUCTION

Smaller device dimensions achieved by technology scaling have enabled a higher level of integration and faster switching speed. However, this trend has increased process variation and power consumption [1]. Process variation results in a significant spread of delay [2]. Furthermore, rapidly-growing power consumption impacts environmental variations such as the supply voltage and the temperature [3]-[6]. Since higher power consumption leads to higher voltage droop and temperature, it is becoming more difficult to achieve higher performance at low supply voltage [3].

Under increasing process and environment variations, it is more demanding for a corner-based design approach to meet aggressive performance specifications [7]. This issue has led to many adaptive design methods that achieve higher performance by overcoming variation factors utilizing replica circuits, which have high correlation with critical path delays of actual circuits [8]-[11]. By tracing the delay of replica circuits, the supply voltage or operating frequency can be dynamically changed to prevent chip failure while maintaining higher performance or lower power consumption. However, different variations in process, voltage, and temperature (PVT) between actual circuits and replica circuits are unavoidable since their geometrical locations are different [4]. Thus, an adaptive design with replica circuits requires a safety margin to prevent chip failure.

An in-situ error detection and correction mechanism is effective method for removing the safety margin while maintaining reliability by managing detected errors [3]-[6]. If an error occurrence is very rare, a system with the in-situ error detection and correction method can achieve higher performance or lower power consumption [4]. However, this approach requires additional power consumption and a performance penalty for recovery or architectural replay [3].

This paper proposes a method of removing the safety margin with a minimized performance penalty. A flip-flop, which allows time borrowing during a time-borrowing window (T_{BW}) and generates a detection signal in the presence of time borrowing, is presented. It also proposes a clock shifter that stretches the clock period by T_{BW} in the detection of time borrowing to pay borrowed time back in the next clock cycle, which can reduce the performance penalty when time borrowing occurs. Therefore, a design employing the time-borrowing flip-flop (TBFF) in conjunction with the clock shifter can reduce power consumption or increase the clock frequency even at a high activation probability of critical paths.

II. PROPOSED METHODOLOGY

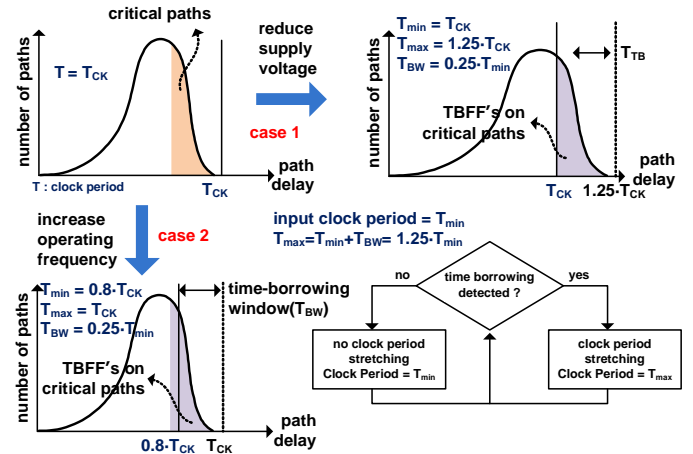


Fig. 1. Concept of the proposed dynamic timing control.

As illustrated in Fig.1, the maximum operating frequency of a circuit is limited by the maximum path delay of critical paths of a circuit. In “case 1” in Fig.1, critical paths whose delay is between T_{CK} and $1.25 \cdot T_{CK}$ will necessarily result in timing failures when the supply voltage is reduced. If the TBFF’s with T_{BW} (T_{BW} is defined by a quarter of the input clock period, T_{min}), are used on critical paths, time slacks which occur at critical paths can be compensated by time borrowing. For the current pipeline stage, time borrowing helps to prevent timing failures. However, it increases the path delay in the next pipeline stage and could end up in a timing failure in the next clock cycle unless the clock period changes. Therefore, by stretching the clock period by T_{BW} in the detection of time borrowing, the possible timing violation in the next pipeline stage can be resolved. As a result, the design with TBFF’s on critical paths and the clock shifter can operate at a clock period of T_{CK} when critical paths are not activated.

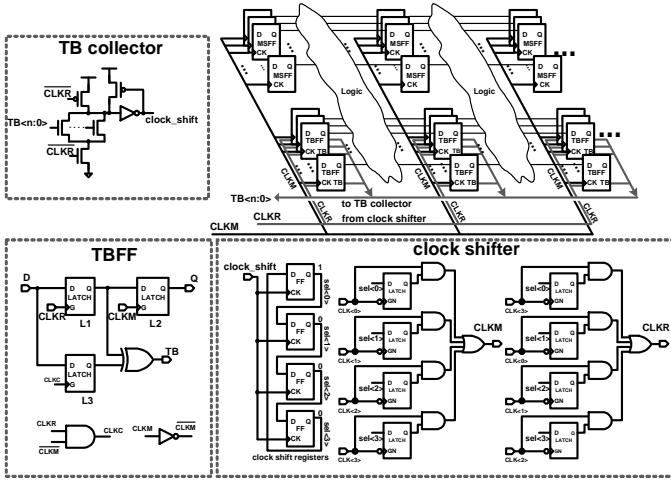


Fig. 2. Pipeline architecture with the proposed method.

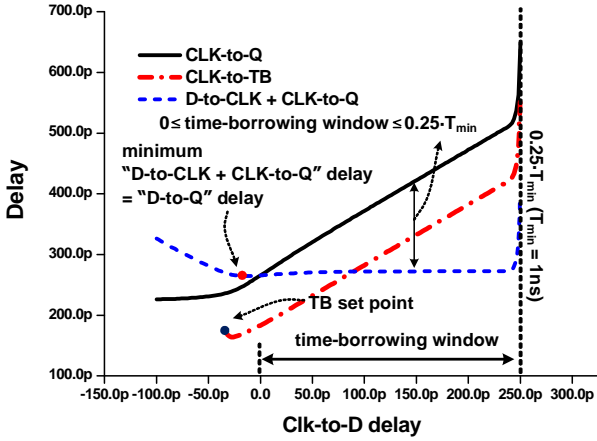


Fig. 3. Simulated TBFF timing.

In the detection of time borrowing, the design will operate dynamically at the clock period of $1.25 \cdot T_{CK}$. In other words, if the activation probability of critical paths is 0, the design with the proposed technique operates at the minimum clock period (T_{min}). On the other hand, if the activation probability of critical paths is 1, the design operates at the maximum clock period (T_{max}). In this way, the design with the proposed scheme can operate at two different clock periods, T_{min} or T_{max} , in response to the activation of critical paths. Since the clock period changes adaptively according to the activation of critical paths, an effective operating frequency (F_{EFF}) is required to estimate the performance. The F_{EFF} for the proposed method is defined as

$$F_{EFF} = (P_c \cdot T_{max} + (1 - P_c) \cdot T_{min})^{-1}, \quad (1)$$

where P_c is the activation probability of critical paths. For the proposed method, the range of F_{EFF} is

$$T_{max}^{-1} \leq F_{EFF} \leq T_{min}^{-1}. \quad (2)$$

Since T_{max} is equal to $1.25 \cdot T_{min}$ for the proposed method, the F_{EFF} is bounded between T_{min} and $1.25 \cdot T_{min}$. In contrast, for the error detection and recovery method proposed in [3]-[6], T_{max} (T_{max} is clock cycles required to recover and replay in the presence of a timing error) is equal to or more than $2 \cdot T_{min}$.

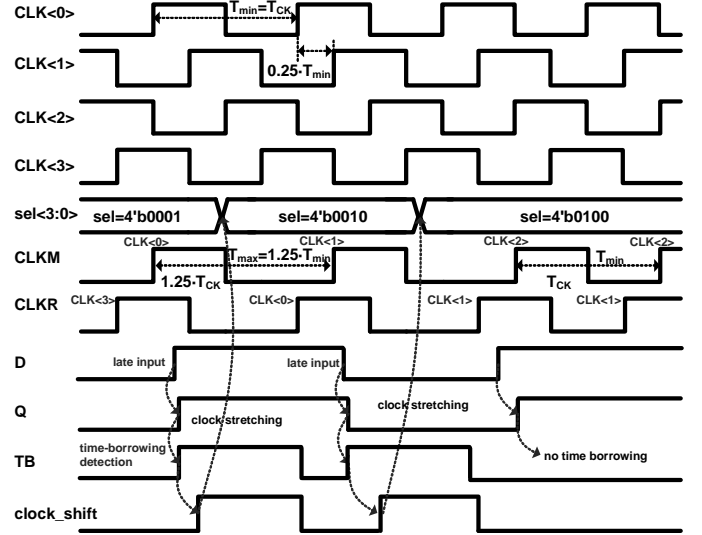


Fig. 4. Conceptual timing diagram of the proposed method.

Therefore, as P_c increases, the performance of the proposed method degrades more slowly than that of the error detection and recovery method.

The same strategy applies to a faster operation when the clock period is reduced by T_{BW} , $0.2 \cdot T_{CK}$, while maintaining the supply voltage, as in “case 2” of Fig.1. As long as path delays of critical paths are within T_{BW} , i.e., the clock stretching range, as illustrated in Fig.1, the T_{min} can be reduced to $0.8 \cdot T_{CK}$ for the faster operation maintaining correct operation. In (1), as the activation probability of critical paths approaches 0, the F_{EFF} can ideally be increased by up to 25%.

III. IMPLEMENTATION

A. Time Borrowing and Detection

The pipeline architecture with the proposed method was implemented as Fig. 2. The TBFF requires a master clock (CLKM) and a reference clock (CLKR) as shown in Fig. 2. The phase of the CLKM is $0.25 \cdot T_{min}$ behind that of the CLKR. The CLKM is used as a global clock for both conventional flip-flops in noncritical paths and TBFF's on critical paths. The CLKR is used to define the time-borrowing window ($T_{BW} = 0.25 \cdot T_{min}$) of the TBFF. Thus, the TBFF can borrow timing up to T_{BW} from the next pipeline stage. When the data of a latch L3 and L1 do not match, a time-borrowing detection signal (TB) indicating time borrowing, is generated to control the clock shifter circuit to change the clock period. However, if the input arrives before the rising edge of the CLKM, the TB is not set and the TBFF behaves like a conventional flip-flop.

The time-borrowing detection is the same concept as the error or transition detection methods studied in [3]-[6]. The main idea of the proposed method is to borrow time from the next pipeline stage and pay the borrowed time back by stretching the clock period at the next clock cycle utilizing the TB. Thus, since the clock period is stretched by T_{BW} , the performance penalty can be minimized.

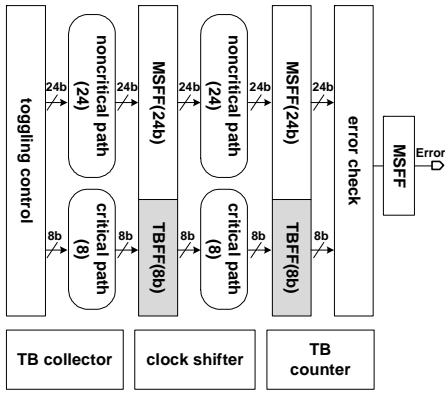


Fig. 5. Prototype for the proposed method.

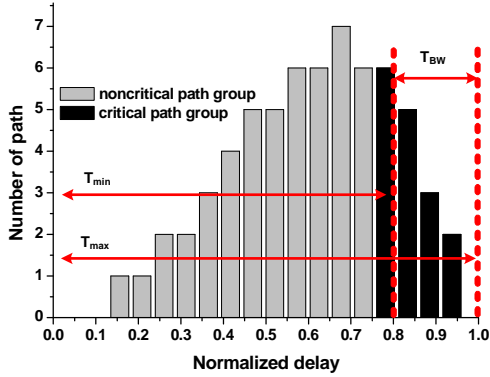


Fig. 6. Normalized delay distribution of delay paths.

The timing characteristics of the TBFF, shown in Fig. 3, depend on the clock period, which defines T_{BW} . To reduce the area and power overhead, latch L3 was designed with smaller transistors than those of the other latches (L1 and L2). Thus, the setup time of latch L3 is larger than that of the other latches (L1 and L2). For this reason, the TB is set ahead of the minimum point of the D-to-Q delay. This behavior helps to prevent failures of time-borrowing detection. If the TB is set after the minimum point of the D-to-Q delay, some time borrowing may not be detected.

Since the TBFF allows time borrowing, it has more stringent requirements for preventing hold time violations. The hold time of the TBFF increases by T_{BW} , $0.25 \cdot T_{min}$. However, the size and power overhead caused by buffer insertion, which prevents the hold time violation, is not critical since TBFF's are used on critical paths. In addition, a design trend towards a shallower depth of pipeline stages [4] relaxes the overhead to fix hold time violations caused by time borrowing.

In architecture-level implementation, a collector block detects the occurrence of any TB from all TBFF's to generate a clock-shift signal that controls the clock shifter as Fig. 2.

B. Clock Stretching

A clock shifter that stretches the clock period dynamically, shown in Fig. 2, is proposed. Whenever the clock shift signal transits from low to high, the clock shift registers change clock

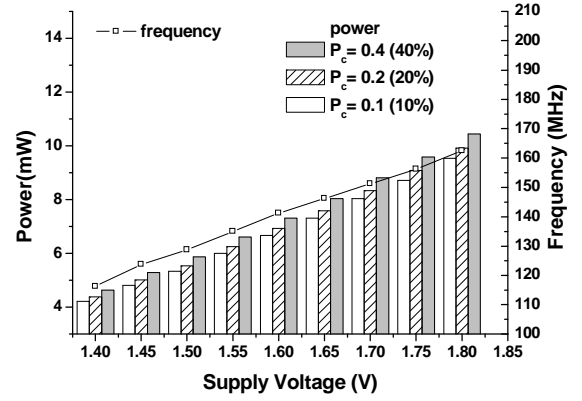


Fig. 7. Measured performance of the design with the conventional method.

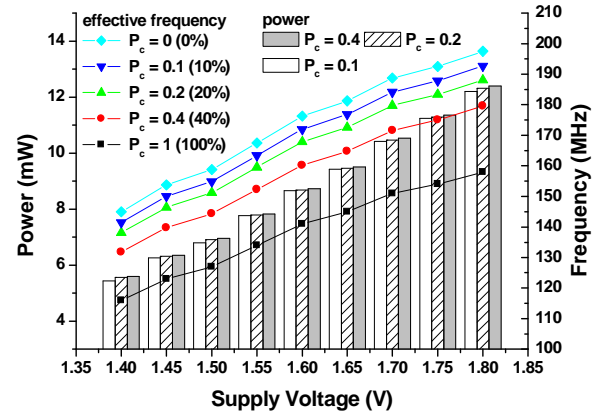


Fig. 8. Measured performance of the design with the proposed method.

phase selection signals (sel), which select one clock phase among four-phase clocks with 90° ($0.25 \cdot T_{min}$) phase differences. When the clock moves from one phase to another, the clock period stretches from T_{min} to T_{max} , as shown in Fig. 4. Therefore, the proposed system with TBFF's and the clock shifter can dynamically operate with two different clock periods in response to the presence of time borrowing. However, assuming that four-phase clocks with the clock period of T_{CK} are provided, a clock period of T_{CK} or $1.25 \cdot T_{CK}$ and T_{BW} of $0.25 \cdot T_{CK}$ can be generated as shown in Fig. 4. In this case, T_{min} and T_{max} are T_{CK} and $1.25 \cdot T_{CK}$, respectively. In the same way, a clock period of $0.8 \cdot T_{CK}$ or T_{CK} and T_{BW} of $0.2 \cdot T_{CK}$ can be generated from four-phase clocks with the clock period of $0.8 \cdot T_{CK}$. Accordingly, T_{min} and T_{max} are $0.8 \cdot T_{CK}$ and T_{CK} , respectively.

As mentioned earlier, the proposed clock shifter requires four-phase clocks. Nevertheless, it can be implemented without additional cost since the four-phase clocks can be provided from a phase-locked loop embedded to provide a system clock.

IV. TEST CHIP

A prototype for the proposed method with a three-stage pipeline was implemented in a 180-nm CMOS (Fig. 5, die-photo in Fig. 10). In the prototype, the activation probabilities

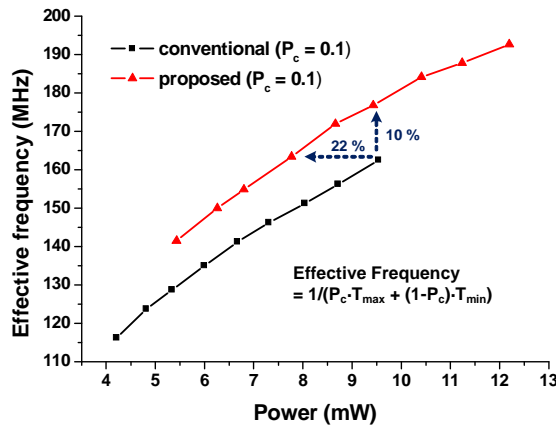


Fig. 9. Measured clock frequency versus power consumption.

of critical and noncritical paths in pipelined stages are controlled separately by programming registers in a toggling control block in Fig. 5. Inverter chains are used to implement different delay paths for the controllability of the activation probability. The normalized path delay distribution of the prototype is shown in Fig. 6. TBFF's are used on 16 critical paths among 64 delay paths. Master-slave flip-flops (MSFF) are used on noncritical paths. For comparison, a reference pipelined system is also implemented with conventional MSFF's in all paths without any dynamic control. A serial peripheral interface bus (SPI) slave block is integrated to program the activation probability of critical paths and noncritical paths. The purpose of a TB counter in Fig. 5 is to count the number of time borrowing occurrences to calculate the F_{EFF} . The count value and the occurrence of an error can be read through the SPI block.

V. MEASUREMENTS

In the chip measurements, the activation probability of critical paths was set to 0.4, 0.2, or 0.1. In the prototype, measured power overhead caused by the TBFF and the clock shifter was 7.9% at the same operating condition as that of the conventional design. In Figs. 7 and 8, the operating frequency was measured changing the input clock frequency until the first error was detected. The measured maximum input frequency for the prototype with the proposed method is when P_c is 0 in Fig. 8. Measured results in different supply voltages show up to a 24.7% increase in the maximum input frequency for the proposed design compared to the conventional design (Figs. 7 and 8) at the same supply voltage. Hence, for a target frequency, the design with the proposed method can operate with a lower power (i.e., at a lower voltage), or for a target power, it can operate at a higher frequency as shown in Fig. 9. For example, considering a 0.1 (10%) activation probability, the proposed design can operate at up to 22% lower power (at the same effective frequency) and up to 10% higher frequency (at the same power) than the conventional design.

VI. CONCLUSION

This paper presented an effective method for preventing timing failures by utilizing time borrowing and clock stretching, thereby eliminating a safety margin. Since an

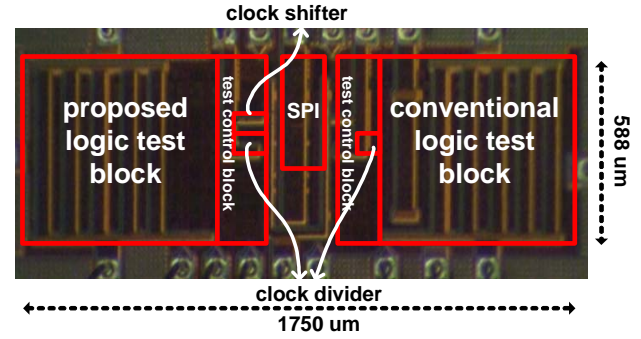


Fig. 10. Die-photo.

additional recovery or replay operation for the error management is not required, the proposed delay-error tolerant method can minimize the performance penalty. As a result, a design with the proposed method can effectively reduce power consumption or increase the operating frequency, which extends the dynamic operating margin of pipelined systems, even at high activation probability of critical paths. A prototype was implemented in a 180-nm CMOS technology and the performance enhancement using the proposed technique was verified.

REFERENCES

- [1] S. Mukhopadhyay, K. Roy, "Modeling and estimation of total leakage current in nano-scaled CMOS devices considering the effect of parameter variation," in *Proc. Int. Symp. Low Power Electronics and Design*, 2003, pp. 172-175.
- [2] Vivek Joshi, David Blaauw, Dennis Sylvester, "Soft-edge flip-flops for improved timing yield: design and optimization," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Nov. 2007, pp. 667-673.
- [3] Keith A et al., "Energy-Efficient and Metastability-Immune Resilient Circuits for Dynamic Variation Tolerance," *IEEE J. Solid-State Circuits*, vol. 44, no. 1, pp. 49-63, Jan. 2009.
- [4] D. Ernst et al., "Razor: A low-power pipeline based on circuit-level timing speculation," in *Proc. IEEE/ACM Int. Symp. Microarchitecture*, Dec. 2003, pp. 7-18.
- [5] S. Das et al., "A Self-Tuning DVS Processor Using Delay-Error Detection and Correction," *IEEE J. Solid-State Circuits*, vol. 41, no. 4, pp.792-804, Apr. 2006.
- [6] S. Das et al., "RazorII: In Situ Error Detection and Correction for PVT and SER Tolerance," *IEEE J. Solid-State Circuits*, vol. 44, no. 1, pp.32-48, Jan. 2009.
- [7] T. Sato and Y. Kunitake, "A Simple Flip-Flop Circuit for Typical-Case Designs for DFM," in *Proc. Int. Symp. Quality Electronic Design*, 2007, pp.539-544.
- [8] J. Tschanz et al., "Adaptive Body Bias for Reducing Impacts of Die-to-Die and Within-Die Parameter Variations on Microprocessor Frequency and Leakage," *IEEE J. Solid-State Circuits*, vol.37, no. 11, pp.1396-1402, Nov. 2002.
- [9] A. K. Uht, "Going beyond Worst-case Specs with TEAtime", *IEEE Computer*, vol.37, no.3, pp 51-56, Mar. 2004.
- [10] J. Tschanz, N.-S. Kim, S. Digne et al., "Adaptive Frequency and Biasing Techniques for Tolerance to Dynamic Temperature-Voltage Variations and Aging," in *IEEE Int. Solid-States Circuits Conf. Dig. Tech. Papers*, Feb. 2007, pp. 292-293.
- [11] A. Drake et al., "A distributed critical-path timing monitor for a 65 nm high-performance microprocessor," in *IEEE Int. Solid-States Circuits Conf. Dig. Tech. Papers*, Feb. 2007, pp. 398-399.