Adaptive Delay Monitoring for Wide Voltage-Range Operation

Jongho Kim*† Gunhee Lee* Kiyoung Choi*

*Department of Electrical and Computer Engineering
Seoul National University
{jongho1119.kim, ppghl88, kchoi}@snu.ac.kr
jongho1119.kim@samsung.com

Yonghwan Kim† Wook Kim† Kyungtae Do† Jungyun Choi†

†Design Technology, System LSI
Samsung Electronics
{yhluz.kim, wook81.kim, kyungtae.do, jungyun74.choi}@samsung.com

***Abstract*— As process technology scales down, circuit delay variations become more and more serious due to manufacturing and environmental variations. The delay variations are hardly predictable and thus require additional design margin and impede the chance to reduce area and power consumption of a chip. One way to alleviate the problem is to measure the circuit delay at run-time and control the supply voltage accordingly through a closed-loop dynamic voltage and frequency scaling (closed-loop DVFS) scheme. The circuit delay is typically measured by a monitoring circuit. However, the key issue of this scheme is the delay mismatch between the monitoring circuit and the target circuit block such as a CPU or a GPU. A large delay mismatch might lose the advantage of closed-loop DVFS. And it becomes worse as the circuit block operates in a wider voltage-range. This paper proposes a novel adaptive delay monitoring scheme for a wide voltage-range operation, which provides a better delay correlation between the monitor and the target compared to conventional monitoring approaches. The proposed approach reduces the average error in the measured delay by up to 45% and the maximum error by up to 68%. The reduction of the error brings the decrease of design margin, resulting in a lower-power and lower-cost design.**

***Keywords*— Monitoring circuit, Delay monitor, Adaptive voltage scaling, Closed-loop dynamic voltage and frequency scaling, Design margin, Wide voltage-range operation**

I. Introduction

Semiconductor products have been perpetually shrinking over the past decades to allow performance enhancement at lower fabrication cost per transistor. However, this process scaling has also brought serious circuit delay variations [1][2][3] mainly due to manufacturing and environmental variations including inter/intra-die variability, temperature shift, supply voltage droop noise, and circuit aging. As CPU/GPU operates in a wide voltage-range (from near-threshold voltage to super-overdrive voltage), circuit delay variations become much worse than previous chip operations [4-8]. Typically, circuit delay variations have been covered by design margins to ensure 'no-error' operations under the variations. This is the most pessimistic approach considering all worst-cases and thus incurs additional costs that would be unnecessary in better-than-worst-cases. Moreover, the problem becomes more serious as the variation becomes larger and larger. Even, it is difficult to determine the optimal design margin considering all the operating conditions at manufacturing test. Especially, this problem might be too serious in CPU/GPU designs because such a block requires a wide-range of operating voltages. Most commercial CPUs/GPUs used to be operated by an open-loop dynamic voltage and frequency scaling (open-loop DVFS) scheme using look-up tables (LUTs) in memory. However, it is recently required to use a closed-loop dynamic voltage and frequency scaling (closed-loop DVFS) scheme which can monitor the circuit delay at run-time. This scheme can prevent from increasing the design margin by circuit delay variation and achieve low-cost, low-area and low-power design.

In the closed-loop DVFS scheme, the key point is how to implement the delay monitoring circuit to estimate the circuit delay accurately. That is, the delay mismatch between the target block (such as a CPU or a GPU) and the monitoring circuit should be minimized, which is directly connected to the effectiveness of the closed-loop DVFS scheme. There have been various circuits proposed to implement a delay monitor [9-19]. They can be classified into two groups according to the dependency on the block or design that the monitor is targeting: generic monitoring circuit and design-dependent monitoring circuit. The generic monitoring circuit [9][10][11] is mainly implemented by a simple inverter-based ring oscillator (RO). It does not have any dependencies on the target block. It is very practical and suitable for a short-time-to-market design because it can be easily implemented and reused for any chip design platforms. It might be difficult, however, to use it for a design implemented by a mix of device types (e.g., multi-threshold voltage design) possibly causing a big gap between the measured delay and the actual delay of the target block. Also, it still requires a large design margin because it typically has a relatively big delay mismatch with the actual critical path of the target block under various operating conditions. Design-dependent monitoring circuit [12-19] is designed to be highly correlated with the target block in terms of delay. However, it typically requires high area overhead and design complexity. And it might be difficult to use it for commercial chip designs because it has longer design turn-around time than the generic one and cannot be reused unless the same target block is reused. The two types of monitoring circuits are presented in the following section in more detail.

In this paper, we propose a novel monitoring circuit and scheme which give both design simplification and accurate

circuit delay estimation. Section II gives an introduction to open-loop DVFS and closed-loop DVFS based on monitoring circuits as well as previous work for monitoring circuits. Section III presents the proposed monitoring circuit and scheme for a wide voltage-range operation and Section IV describes the design methodology flow. Section V presents the experimental results and the analyses. Finally, Section VI summarizes our work with concluding remarks.

II. BACKGROUND

Dynamic voltage and frequency scaling (DVFS) is widely used for reducing power consumption during off-peak processing time and preventing from thermal overheating problems. The goal of this technique is to run the chip at the lowest possible voltage while achieving the desired operating frequency. According to the use of feedback information from delay monitoring circuits, the DVFS approaches can be classified into two different schemes: open-loop DVFS and closed-loop DVFS. Following subsections briefly introduce the two DVFS schemes as well as various monitoring circuits used for the closed-loop DVFS scheme.

A. Open-loop DVFS scheme

Open-loop DVFS is the most commonly used in DVFS scheme. The operating voltage for each desired operating frequency can be determined at the manufacturing test step while considering operating temperature conditions. Then the frequency-to-voltage mapping information is generally stored in a look-up table (LUT) and used by the power management controller to scale the supply voltage up and down as requested by the application. Typically a large design margin is assigned to the operating voltage for each target frequency stored in the LUT for safety reason. Thus, there is a limitation in reducing power consumption aggressively, since the operating voltage is pre-determined and cannot be adjusted dynamically depending on the run-time conditions.

B. Closed-loop DVFS scheme

Closed-loop DVFS is recently emerged to get more aggressive power reductions in ultra-low power competitive markets. The design margin in the supply voltage should be reduced as much as possible. Since the actual operating speed changes with various PVT (Process/Voltage/Temperature) conditions, there is a need to calculate and assign the optimal voltage dynamically at run-time. For example, circuit aging, supply voltage droop noise, and temperature shift always occur in the real environment and they affect the actual chip operating speed. This problem can be resolved by a feedback loop based on a delay monitoring circuit which provides information on how fast or slow the chip is actually running. As shown in Fig. 1, the feedback loop is facilitated by a monitoring circuit that enables closed-loop DVFS, where the operating voltage is adaptively scaled to an optimal point.

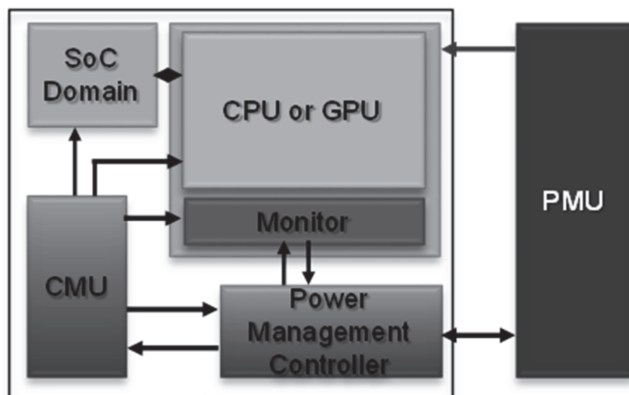


Fig. 1. Closed-loop DVFS system.

Compared to open-loop DVFS, closed-loop DVFS does not require a large design margin because it tracks the operating speed at run-time. However, it is impossible to remove the whole design margin because there can be a delay mismatch between the target block and the monitoring circuit. Therefore, the advantage of closed-loop DVFS depends on the delay correlation between the target block and the monitoring circuit, and thus it is the most important factor to decrease the delay mismatch between the two circuits in such a system. In the next section, we introduce previous work on monitoring circuits.

C. Previous work on monitoring circuits

The accuracy of delay monitoring circuit is very important to maximize the effect of closed-loop DVFS. There have been many researches on monitoring delays and various circuits have been proposed. However, the circuits can be simply classified into two categories: generic monitoring circuit and design-dependent monitoring circuit. A generic monitoring circuit is typically designed as a simple inverter-based ring oscillator (RO). Process-specific ROs (PSROs) have been proposed to measure process parameters or variations of a chip [9][10]. Phase-locked loop (PLL) is used for an alternative monitoring circuit [11]. This monitoring circuit is very simple and easy to design. Also, it does not generate large area overhead and can be easily reused for any other chip designs. But, it is in general less accurate than a design-dependent monitoring circuit and thus incurs a large design margin due to the large delay mismatch. A design-dependent monitoring circuit is tuned so that its delay characteristics are better correlated with those of the target block. Thus it is more accurate than generic PSROs, but many calibrations and parameter storage resources are required. Such a design-dependent delay monitor can be implemented based on a design-specific delay model [12][13]. A design-dependent RO can be synthesized according to the target design and process information [14]. It is relatively simple and has lower area overhead than other kinds of design-dependent monitor, but additional design turnaround time and characterization of the target block are required. Also, it cannot be reused for other chip designs. Critical path replica and in-situ monitors which can analyze the critical path of the target block have been presented [15][16][17]. Although it has good correlation with the target block, it generates large area overhead and long design turnaround time. Reconfigurable monitors presented in [18] can be flexibly tuned according to the circuit delay of the target block. Hence, it provides more accurate circuit delay

estimation, but still has complex calibration and large area overhead. Delay sensor is integrated directly into the critical paths to measure process/environmental variations [19]. It provides the most accurate circuit delay information because it directly tracks the real critical path, but it incurs higher design complexity compared to other monitoring circuits.

In the next section, we present the proposed monitoring circuit and scheme, which is based on a generic monitoring circuit but has accuracy comparable to that of design-dependent monitoring circuits.

III. PROPOSED CIRCUIT AND SCHEME

A. Conventional approach with generic monitoring circuits

Generic monitoring circuits are widely used for commercial chip designs requiring short time-to-market due to the simplicity and reusability. It is typically implemented by a single device type, while the target block such as a CPU or a GPU tends to be implemented by a mix of multiple device types supported by a multi-threshold-voltage library for achieving both low power and high performance at the same time. Such a design with multiple device types exhibits different delay responses in different PVT (especially voltage) conditions. Besides, the operating voltage range of a CPU/GPU becomes wider, from near-threshold voltage to super-overdrive voltage. As a result, the delay mismatch is accelerated requiring a large design margin because the design margin is dominated by the maximum delay mismatch. That is, generic monitoring circuit has the limitation by using only one device type among the types from a multi-threshold-voltage library. Examples include, low-threshold-voltage (LVT) transistors, regular-threshold-voltage (RVT) transistors and high-threshold-voltage (HVT) transistors. Alternatively, it is possible to mix these types with a specific ratio. However, it's very difficult to determine the optimal ratio of different threshold-voltage transistors from many critical paths of the target block. Thus, generic monitors mainly choose LVT under the assumption that most of critical paths consist of LVT transistors or choose HVT when high delay sensitivity is needed. Any of such implementation decisions still has a limitation in estimating the circuit delay of the target block implemented with multi-threshold-voltage transistors.

B. Proposed monitoring circuit

To overcome the limitation of a single generic monitoring circuit, we develop a combination of three generic monitoring circuits implemented respectively with LVT, RVT and HVT transistors, which are used for the target block implementation. We illustrate a simple view of the proposed monitoring circuit structure in Fig. 2. Each chain has different delay characteristics according to PVT (especially voltage) conditions. It is possible to add an additional chain that has different delay characteristics, for example, with a mix of multi-threshold voltage transistors or different gate types. In our experiments, we implement the monitoring circuits with 2-input NAND and 2-input NOR cells because INV cells are not good for control (it has only one-input). Also, we use NAND and NOR cells with one-to-one ratio, in order to have the fair effects of PMOS and NMOS transistors to circuit delay

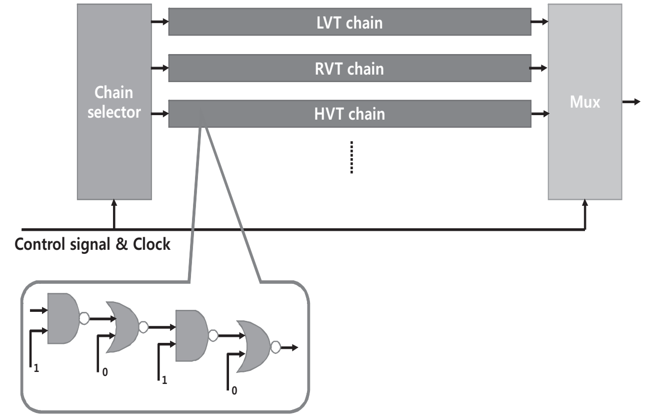


Fig. 2. Proposed monitoring circuit structure.

estimation. The delay chain of NAND is dominated by NMOS transistor delay, while the delay chain of NOR is by PMOS transistor delay and two cases might cause problems to estimate circuit delay accurately in process skew corners (e.g., SF/FS process corners). This monitoring circuit outputs information on the number of cell stages in the selected chain, through which the input signal propagates within one-clock cycle. The number becomes large if the supply voltage is scaled up and becomes small if voltage is scaled down. Thus the cell delay is easily known through the output number.

To make the effect of PVT variation on the monitoring circuit as close as that of the target block, we propose to place the monitoring circuit closely to the target block. Also, it is desirable to share the same supply voltage rail to consider the supply voltage droop noise of the target block.

C. Proposed monitoring scheme

Critical paths of the target block and the delay chains have different delay-voltage characteristics throughout the whole operating voltage range (see Fig. 4) because they have different ratio of LVT/RVT/HVT cells and different gate types in their own paths. Typically, the delay-voltage characteristics of HVT are the most sensitive, while those of LVT are the least sensitive.

We consider using different delay chains as the voltage and/or temperature change. As the supply voltage or temperature changes, the delay also changes; the rate of the delay change of the target block may match well with that of a

Proposed Algorithm

Input : *voltage range index set* $\mathbf{V} = [v_1\ v_2\ v_3\ \ldots\ v_n]^T$

chain index set $\mathbf{C} = [c_1\ c_2\ c_3\ c_4\ \ldots\ c_m]^T$

Output : *chain mapping set* $\mathbf{P_t} = [p_1\ p_2\ p_3\ \ldots\ p_n]$, $p_n = [x_{n,1}\ x_{n,2}\ x_{n,3}\ \ldots\ x_{n,m}]^T$, $x_{n,m} = 0 \text{ or } 1$

For all v_n,

Calculate the delay-voltage slope of critical path and all chains $(slope_{critical}, slope_{c_1}, \ldots, slope_{c_m})$

Find $\text{Min}\ \{|slope_{critical} - slope_{c_1}|, |slope_{critical} - slope_{c_2}|, \ldots, |slope_{critical} - slope_{c_m}|\}$

If $|slope_{critical} - slope_{c_k}|$ *is minimum,*

$p_n = [x_{n,1}\ x_{n,2}\ x_{n,3}\ \ldots\ x_{n,m}]^T \rightarrow \text{all } x_{n,m} = 0 \text{ except } x_{n,k} = 1$

End

Algorithm 1. Chain mapping algorithm for each voltage range.

delay chain. At different voltage and temperature, however, a different chain may match better, and that is why we use multiple delay chains.

Among the multiple chains, the proposed monitoring scheme adaptively selects a chain that matches best with the target block at the current voltage/temperature during run-time. The voltage steps are determined by the resolution of the PMIC (Power Management IC) or the designer's decision. The finer the voltage steps are, the accuracy becomes higher. For this, critical paths of the target block and the delay chains should be characterized to obtain delay-voltage slope throughout the entire operating voltage range. Regarding the temperature, since the delay is much less sensitive to temperature variations[1] , the entire range is divided into only three sub-ranges. Then, for each voltage step and temperature sub-range, the chain having the most similar delay-voltage characteristic with the critical path is mapped. That is, the proposed scheme selects the chain that has delay-voltage slope closest to that of the critical path and use it for circuit delay estimation at the voltage and temperature. The mapping information is stored in an LUT and the power management controller selects a proper chain for the current voltage and temperature by using the LUT.

To explain how we determine the contents of the LUT, suppose there are m chains, n operating voltage steps. We define a vector $\boldsymbol{V} = [v_1\ v_2\ v_3\ \dots\ v_n]^T$ of n voltage steps (for i-th voltage step, v_i=1 and all other entries are 0's) and a vector $\boldsymbol{C} = [c_1\ c_2\ c_3\ c_4\ \dots\ c_m]^T$ of m chains (for j-th chain, c_j=1 and all other entries are 0's). Then, for each temperature range t, we determine an $m \times n$ matrix $\boldsymbol{P_t} = [p_1\ p_2\ p_3\ \dots\ p_n]$ that maps a voltage step to a chain as follows.

$$\boldsymbol{C} = \boldsymbol{P_t} \cdot \boldsymbol{V} \qquad (1)$$

where each column p_i of $\boldsymbol{P_t}$ has all 0's except for one entry which is 1. Thus once the current voltage step is determined, then the proper chain is determined by $\boldsymbol{P_t}$. Algorithm 1 shows how we determine the entries of matrix $\boldsymbol{P_t}$. Once the matrix is determined, it becomes the contents of the LUT. This approach gives much better delay correlation between the monitoring circuit and the target block than existing generic monitoring circuits without incurring large area overhead and design complexity. In section VI, we analyze the experimental results in detail.

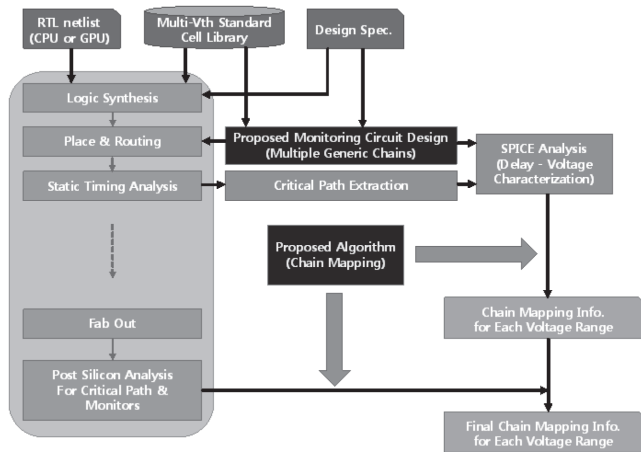


Fig. 3. Design methodology of the proposed approach.

[1] In the range of near threshold voltage, the delay may be sensitive to temperature variations, but it is not the case in the voltage range of our interest in this paper

IV. Design Methodology of Proposed Approach

The design methodology of the proposed approach is shown in Fig. 3. It is very easy to plug the approach into an existing design flow. First, we implement the monitoring circuit with multiple chains following the design specifications such as number of chains using a multi-threshold voltage library. Then, we extract the critical paths of the target block and analyze the delay-voltage characteristics of both the critical paths and the multiple chains under various PVT conditions. Based on the analysis, the proposed algorithm extracts the chain mapping information for each voltage step. After this analysis, an additional chain can be added to the monitoring circuit for better delay correlation. This mapping information might be refined at the post silicon step, based on the physical chip test. It does not generate the additional test cost because the chips should be tested anyway for each frequency level to find the optimal voltage. During this test, it is required to read the output value of each chain for each voltage step. The final mapping information is stored in the LUTs and power management controller uses it for circuit delay estimation at run-time.

V. Experimental Result

To verify our proposed monitoring circuit and scheme, we use ARM Cortex-A53 implemented by Samsung 14nm FinFET technology. It's implemented by three multi-threshold voltage libraries. Following the proposed design methodology in Fig. 3,

Table I. Experimental Environments

Target Design	ARMv8-A Cortex-A53
Process Technology	Samsung 14nm FinFET
Library	Multi-Vth library (3-types)
Process Corner	NN
Operating Voltage [V]	0.6 - 1.2
Temperature [℃]	-40 / 25 / 125

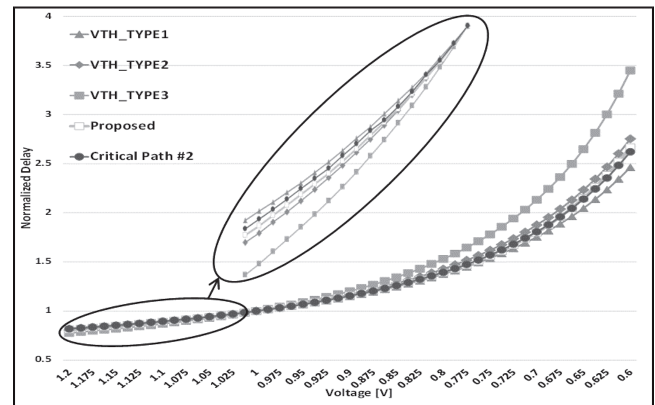
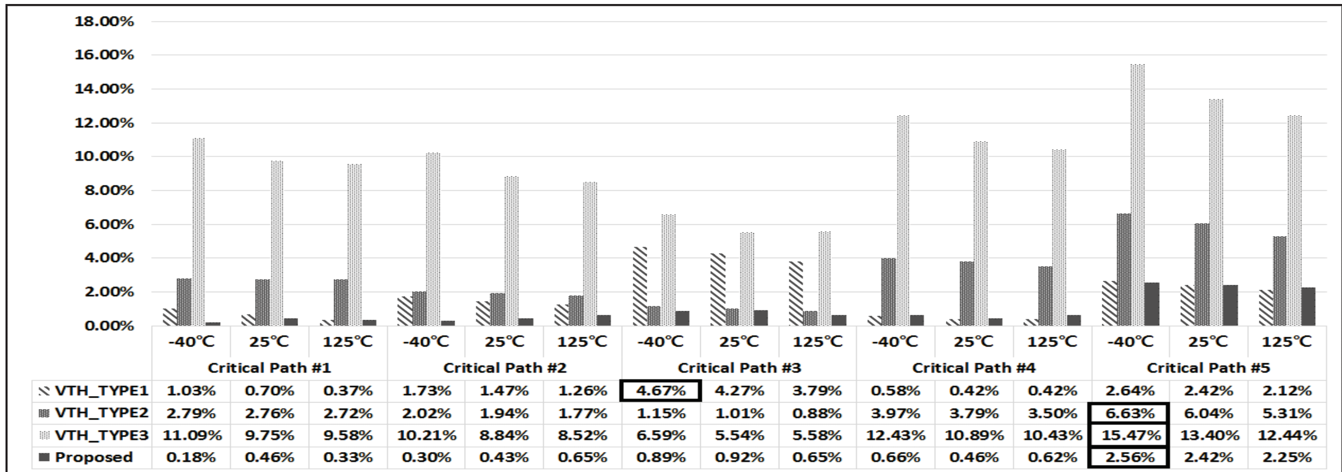
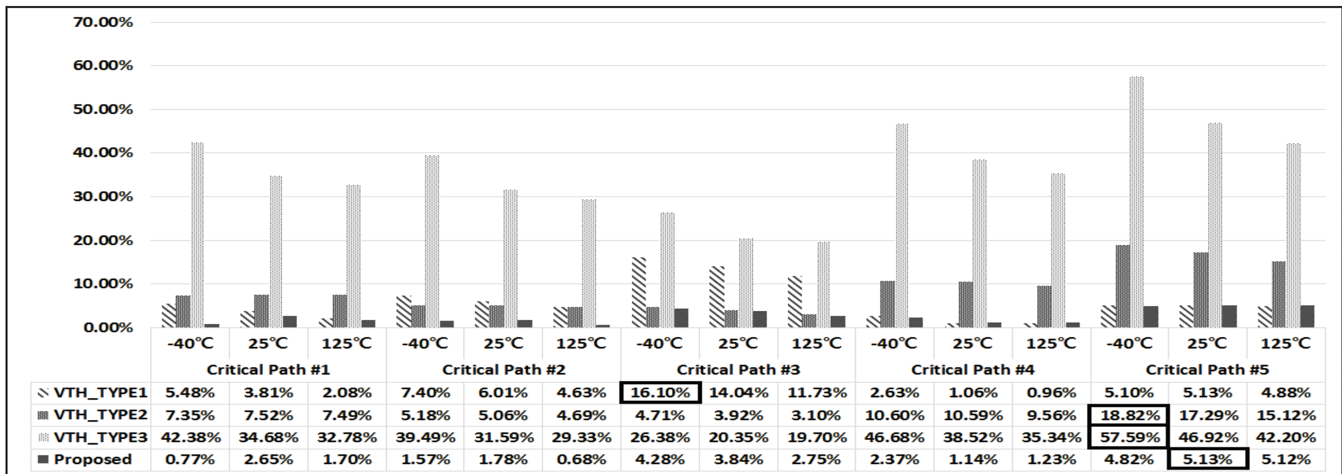


Fig. 4. Delay-voltage characteristic curve of critical path #2 for each chain and the proposed approach at 25℃.



	Critical Path #1			Critical Path #2			Critical Path #3			Critical Path #4			Critical Path #5		
	-40℃	25℃	125℃	-40℃	25℃	125℃	-40℃	25℃	125℃	-40℃	25℃	125℃	-40℃	25℃	125℃
VTH_TYPE1	1.03%	0.70%	0.37%	1.73%	1.47%	1.26%	4.67%	4.27%	3.79%	0.58%	0.42%	0.42%	2.64%	2.42%	2.12%
VTH_TYPE2	2.79%	2.76%	2.72%	2.02%	1.94%	1.77%	1.15%	1.01%	0.88%	3.97%	3.79%	3.50%	6.63%	6.04%	5.31%
VTH_TYPE3	11.09%	9.75%	9.58%	10.21%	8.84%	8.52%	6.59%	5.54%	5.58%	12.43%	10.89%	10.43%	15.47%	13.40%	12.44%
Proposed	0.18%	0.46%	0.33%	0.30%	0.43%	0.65%	0.89%	0.92%	0.65%	0.66%	0.46%	0.62%	2.56%	2.42%	2.25%

Fig. 5. Average errors in the delay estimation by monitoring circuits.



	Critical Path #1			Critical Path #2			Critical Path #3			Critical Path #4			Critical Path #5		
	-40℃	25℃	125℃	-40℃	25℃	125℃	-40℃	25℃	125℃	-40℃	25℃	125℃	-40℃	25℃	125℃
VTH_TYPE1	5.48%	3.81%	2.08%	7.40%	6.01%	4.63%	16.10%	14.04%	11.73%	2.63%	1.06%	0.96%	5.10%	5.13%	4.88%
VTH_TYPE2	7.35%	7.52%	7.49%	5.18%	5.06%	4.69%	4.71%	3.92%	3.10%	10.60%	10.59%	9.56%	18.82%	17.29%	15.12%
VTH_TYPE3	42.38%	34.68%	32.78%	39.49%	31.59%	29.33%	26.38%	20.35%	19.70%	46.68%	38.52%	35.34%	57.59%	46.92%	42.20%
Proposed	0.77%	2.65%	1.70%	1.57%	1.78%	0.68%	4.28%	3.84%	2.75%	2.37%	1.14%	1.23%	4.82%	5.13%	5.12%

Fig. 6. Maximum errors in the delay estimation by monitoring circuits.

we extract 20 critical paths and implement three generic chains with corresponding multi-threshold voltage libraries (VTH_TYPE1/VTH_TYPE2/VTH_TYPE3). To analyze the delay-voltage characteristics of the critical paths and the chains, we run HSPICE under various PVT conditions. We summarize the experimental environments in Table I.

We choose the representative five paths among the 20 extracted critical paths and analyze the delay-voltage characteristics of the five paths and the chains while sweeping voltage from 0.6V to 1.2V with 12.5mV granularity. In Fig. 4, we draw the delay-characteristic curve of a critical path, chains, and the proposed approach under a given specific condition (e.g., 25 ℃) and normalize the delay to that at 1.0V. VTH_TYPE3 chain is most sensitive to voltage changes and then VTH_TYPE2 and VTH_TYPE1 come in that order. That is, the delay-voltage slope of VTH_TYPE3 is largest and the slope of VTH_TYPE1 is smallest. And we confirm that the delay is increasing considerably faster at low voltage region. As can be seen from the figure, the delay-voltage characteristic obtained by our proposed approach is closest to that of critical path.

Fig. 5 and Fig. 6 show the comparison results of our proposed approach for various experimental conditions. Fig. 5 shows average errors in the delay estimation by the monitoring circuits. In critical paths #1, #2 and #3, the error rate of the proposed approach is much smaller than that of other single generic monitoring circuits. Compared to the best result of single generic monitoring circuit, the error rate decreases by about 10% ~ 80%. In critical paths #4 and #5, the error rate of the proposed approach is comparable with the best result of the single generic monitoring circuits. Some results are same and other results are slightly better or worse, but it shows almost the same error rate. Fig. 6 shows the maximum errors in the delay estimation. In the view point of design margin reduction, maximum error rate is much more important than the average error rate because design margin is determined to ensure 'no-error' operation against the worst case. Similarly to the result of average error rates, the proposed approach shows better error rates in critical paths #1, #2 and #3 than other single generic monitoring circuits. In case of critical paths #4 and #5, it shows error rates similar to those of the best result of single generic monitoring circuits.

As described above, if we consider each critical path separately, the proposed approach gives similar or about 10% ~ 80% improved delay correlation. However, the target block should be designed considering all critical paths. In other words, all critical paths should be fully considered to determine the design margin. Thus the design margin should be determined by the critical path that renders the worst delay estimation error. In our experiment, the thick-lined boxes of Fig. 5 and Fig. 6 show the worst results as well as the corresponding critical paths. In case of average error rate, the worst case of VTH_TYPE1 chain is 4.67% on critical path #3, that of VTH_TYPE2 chain is 6.63% on critical path #5, and that of VTH_TYPE3 chain is 15.47% on critical path #5. Compared to the best result of single monitoring circuit (VTH_TYPE1 chain), the proposed approach reduces the error rate by up to 45% (the difference is much larger than that obtained when considering only one critical path such as #5). Moreover, considering that the design should actually consider the maximum error rate, we should use the data for maximum error rate, which shows error rate reduction by up to 68% (from 16.10% down to 5.13%). The improvement of this delay estimation error rate brings the decrease of design margin and lowers design cost and power consumption.

VI. CONCLUSION

In this paper, we propose a monitoring circuit composed of multiple generic chains and a method to adaptively select a proper chain for a wide voltage-range operation. The proposed design methodology is easy to plug into existing design flows. Also, it reduces errors due to delay mismatch by 45% on average and maximum error by up to 68% without large area overhead and cost. It can significantly reduce the design margin for compensating the delay mismatch between the monitoring circuit and the target block in the closed-loop DVFS scheme. In an advanced process technology, the effect of variations on circuit delay can be much more serious and thus the accuracy improvement of measured circuit delay becomes very important.

ACKNOWLEDGMENT

This work was supported by System LSI and Samsung Advanced Institute of Technology, Samsung Electronics Co., Ltd.

REFERENCES

[1] G. F. Taylor, "Where are we going? Product scaling in the system on chip era," in *Proceedings of the IEEE International Electron Devices Meeting (IEDM)*, pp. 17.1.1-17.1.3, Dec. 2013.

[2] K. J. Kuhn, "Reducing variation in advanced logic technologies: Approaches to process and design for manufacturability of nanoscale CMOS," in *Proceedings of the IEEE International Electron Devices Meeting (IEDM)*, pp. 471-474, Dec. 2007.

[3] K. J. Kuhn, M. D. Giles, D. Becher, P. Kolar, A. Kornfeld, R. Kotlyar, S. T. Ma, A. Maheshwari, and S. Mudanai, "Process technology variation," *IEEE Transactions on Electron Devices*, vol. 58, no. 8, pp. 2197-2208, Aug. 2011.

[4] S. Jain, S. Khare, S. Yada, V. Ambili, P. Salihundam, S. Ramani, S. Muthukumar, M. Srinivasan, A. Kumar, S. K. Gb, R. Ramanarayanan, V. Erraguntla, J. Howard, S. Vangal, S. Dighe, G. Ruhl, P. Aseron, H. Wilson, N. Borkar, V. De, and S. Borkar, "A 280mV-to-1.2V wide-operating-range IA-32 processor in 32nm CMOS," in *IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, pp. 66-68, Feb. 2012.

[5] T. Tekeste, A.Shabra, D. Boning, and I. Elfadel, "Variability analysis of a 28nm near-threshold synchronous voltage converter," in *Proceedings of the IEEE International Conference on Electronics, Circuits and Systems (ICECS),* pp. 723-726, Dec. 2013.

[6] U. R. Karpuzcu, N. S. Kim, and J. Torrellas, "Coping with parametric variation at near-threshold voltages," *IEEE Micro*, vol. 33, no. 4, pp. 6-14, Jul.-Aug. 2013.

[7] S. Seo, R. G. Dreslinski, M. Who, Y. Park, C. Charkrabari, S. Mahlke, D. Blaauw, and T. Mudge, "Process variation in near-threshold wide SIMD architectures," in *Proceedings of the ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp. 980-987, Jun. 2012.

[8] M. Seok, G. Chen, S. Hanson, M. Wieckowski, D. Blaauw, and D. Sylvester, "CAS-FEST 2010: Mitigating variability in near-threshold computing," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 1, pp. 42-49, Mar. 2011.

[9] M. Bhushan, A. Gattiker, M. B. Ketchen, and K. K. Das, "Ring oscillators for CMOS process tuning and variability control," *IEEE Transactions on Semiconductor Manufacturing*, vol. 19, no. 1, pp. 10-18, Feb. 2006.

[10] L. T.-N. Wang, N. Xu. S. Toh, A. R. Neureuther, T.-J. K. Liu, and B. Nikolic, "Parameter-specific ring oscillator for process monitoring at the 45 nm node," in *Proceedings of the IEEE Custom Integrated Circuits Conference (CICC)*, pp. 1-4, Sep. 2010.

[11] K. Kang, S. P. Park, K. Kim, and K. Roy, "On-chip variability sensor using phase-locked loop for detecting and correcting parametric timing failures," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 18, no. 2, pp. 270-280, Feb. 2010.

[12] T.-B. Chan, A. Pant, L. Cheng, and P. Gupta, "Design dependent process monitoring for back-end manufacturing cost reduction," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD),* pp. 116-122, Nov. 2010.

[13] N. E. Lustig, K. L. Saenger, and H.-M. Tong, "In-situ endpoint detection and process monitoring method and apparatus for chemical-mechanical polishing," U.S. Patent No. 5,433,651. Jul. 1995.

[14] T.-B. Chan, P. Gupta, A. B. Kahng, and L. Lai, "DDRO: A novel performance monitoring methodology based on design-dependent ring oscillators," in *Proceedings of the International Symposium on Quality Electronic Design (ISQED)*, pp. 633-640, Mar. 2012.

[15] K. Seno, S. Takahiro, and T. Kondo, "Semiconductor device replica circuit for monitoring critical path and construction method of the same," U.S. Patent No. 6,414,527, Jul. 2002.

[16] D. Fick, N. Liu, F. Zhiyoong, M. Fojtik, J.-S. Seo, D. Sylvester, and D. Blaauw, "In situ delay-slack monitor for high-performance processors using an all-digital self-calibrating 5ps resolution time-to-digital converter," in *IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, pp. 188-189, Feb. 2010.

[17] X. Wang, M. Tehranipoor, and R. Datta, "Path-RO: A novel on-chip critical path delay measurement under process variations," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD),* pp. 640-646, Nov. 2008.

[18] A. J. Drake, R. M. Senger, H. Singh, G. D. Carpenter, and N. K. James, "Dynamic measurement of critical-path timing," in *Proceedings of the IEEE International Conference on Integrated Circuit Design and Technology and Tutorial (ICICDT)*, pp. 249-252, Jun. 2008.

[19] X. Wang, M. Tehranipoor, S. George, D. Tran, and L. Winemberg, "Design and analysis of a delay sensor applicable to process/environmental variations and aging measurements," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 20, no. 8, pp. 1405-1418, Aug. 2012.