Optimal Zigzag (OZ): An Effective yet Feasible Power-Gating Scheme Achieving Two Orders of Magnitude Lower Standby Leakage

Kyu-won Choi, Yingxue Xu and Takayasu Sakurai
Center for Collaborative Research, and Institute of Industrial Science, University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo, Japan, 153-8505
Phone: +81-3-5452-6253, Fax: +81-3-5452-6252, {kwchoi,xyx,tsakurai}@iis.u-tokyo.ac.jp

***Abstract*— An effective yet easy-to-implement power-gating scheme is proposed to reduce leakage by two orders of magnitude. Proposed Optimal Zigzag ("OZ scheme") utilizes optimal combination of input-phase-forcing (IPF) and Zigzag cut-off CMOS on critical paths, while uses high-V_{TH} cells on non-critical paths. Various low-leakage design alternatives including dual-V_{TH}, off-off stacking, MTCMOS, and selective MTCMOS are compared with the proposed OZ scheme in terms of delay, leakage power, and area. With 65-nanometer technology and ten typical benchmark circuits, the OZ scheme has been shown to achieve 70% less leakage, 9.4% faster speed, and 37% smaller area than selective MTCMOS which is the most widely used power-gating approach to date.**

I. Introduction

The power dissipation, especially leakage power, continues to increase as technology scales. Several techniques have been proposed to reduce leakage power such as dual-V_{TH} [1], multi-threshold (MT) CMOS [2], off-off MOS stacking [3], selective MTCMOS (SMT) [4], and Zigzag super cut-off CMOS (ZSCCMOS) [5]. Frequently used dual-V_{TH} scheme is shown in Fig.1(a). The dual-V_{TH} scheme has eventually no area overhead and can reduce leakage current in both active mode and sleep mode. The transistors on critical paths, however, are still leaky due to the use of low-V_{TH} transistors and it is difficult to reduce the leakage current further by the dual-V_{TH} technique alone.

On the other hand, the original MTCMOS, which is one of the early power-gating schemes, has a power switch size problem [6]. The power switch is shared by many gates and the estimation of the delay of a power-gated logic block becomes very difficult since the delay depends on the input vector.

In order to overcome the shortcomings of the dual-V_{TH} and MTCMOS, selective MTCMOS (SMT) scheme as shown in Fig.1(b) has been proposed and rapidly accepted in product-oriented designs in the industry [7]. In the SMT, power-gated cells are selectively applied to low-V_{TH} cells in dual-V_{TH} design. Thus, the remaining leakage paths in the dual-V_{TH} design can be cut off by the high-V_{TH} power switch in the cell in the standby mode. In the SMT, the power switch is inserted in each gate and consequently the characterization of the cell delay is straight-forward and the SMT-based design is easy and practical using the existing design tools.

A drawback of the SMT, however, is the keeper insertion in each of the power-gated cell (MT cell). When the power switch is cut off in the MT cell, the output of the MT cell can be in Hi-Z state and can be any voltage which in turn gives rise to high leakage in the non-MT cell connected to the output of the MT cell. In order to overcome this problem, additional keeper circuit is needed for MT cells and this results in large area overhead and delay overhead.

To solve this problem, the OZ scheme shown in Fig.1(c) is proposed in this paper which applies Zigzag cut-off CMOS selectively as power gating after applying dual-V_{TH} scheme and optimal input-phase-forcing (IPF) to maximize the number of off-off stacked cells in the critical path. This reduces the area and delay overhead of SMT while maintaining the leakage as low.

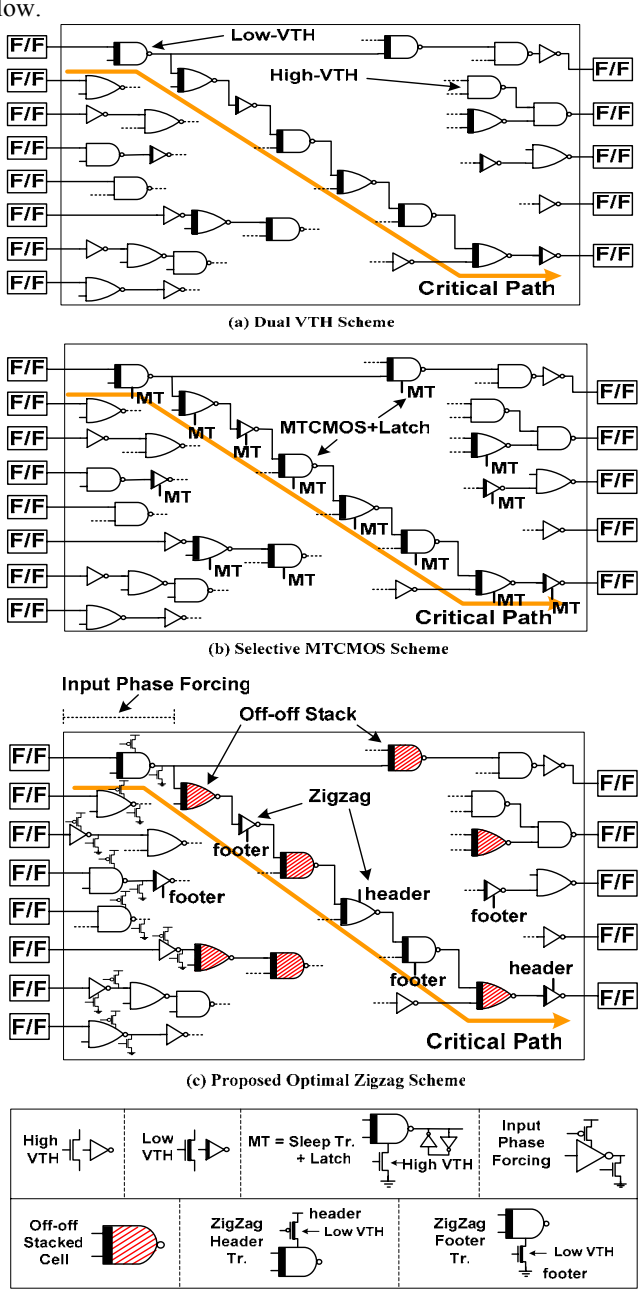
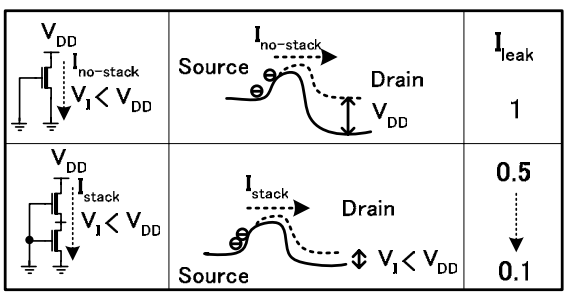


Figure 1. Conceptual diagram

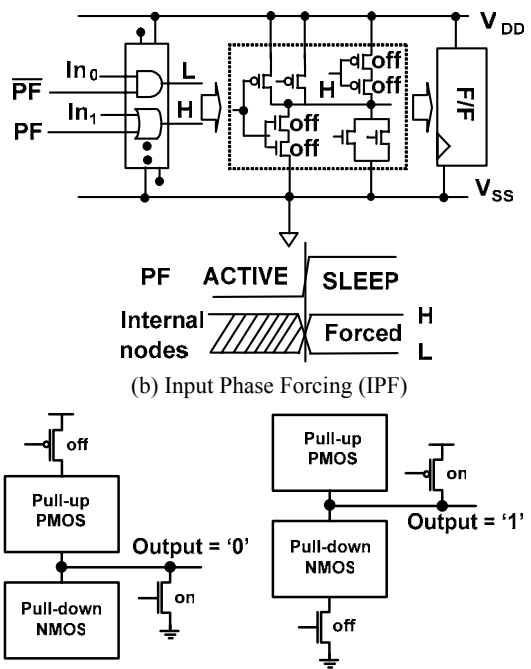
II. Optimal Zigzag (OZ) Technique

In order to explain the proposed Optimal Zigzag scheme, two existing low-leakage schemes should be described first. One of the schemes to reduce leakage is the stacking of off-state transistors, or off-off stacking. Thanks to the Drain Induced Barrier Lowering (DIBL) effect, the off-off transistor stacked structure significantly reduces leakage as shown in Fig. 2. By carefully choosing a primary input vector (Input Phase Forcing: IPF), one can maximize the number of off-off stacked gates in a circuit. For example, if both input of 2NAND are '0', the 2NAND shows one order of magnitude less leakage than '0' and '1' input case. Random search techniques have been used to find the best input combinations [3].

Instead of applying a specific input vector, first gates are modified to mimic the specific input vector as shown in Fig.2(b). It is always possible to modify the first gate by adding two transistors to achieve the required output of the gate as is shown in Fig.2(c).



(a) Off-off stacking effects



(b) Input Phase Forcing (IPF)

(c) Circuit to achieve input phase forcing

Figure 2. Off-off stacking

One more technique that should be explained is the Zigzag super cut-off CMOS (ZSCCMOS) power-gating scheme. In the ZSCCMOS, before entering into a standby mode the input phase forcing is applied to keep each internal node at '0' or '1' state as shown in Fig.3. After the phase forcing, no matter how long the standby time continues, the internal states do not change, because for example, the gate with output '0' is directly connected to V_{SS}. Thus there is no need to insert the keeper being different from the case of the MT cell explained above.

In the ZSCCMOS, the gate of a power switch is over-driven to reduce the leakage. For example, a V_N (negative voltage) generator for NMOS can be implemented without overstress to transistor gate oxide as shown in Fig.4 which was proposed in [5].

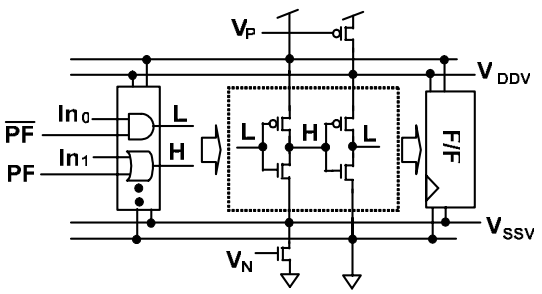


Figure 3. Zigzag super cut-off CMOS (ZSCCMOS)

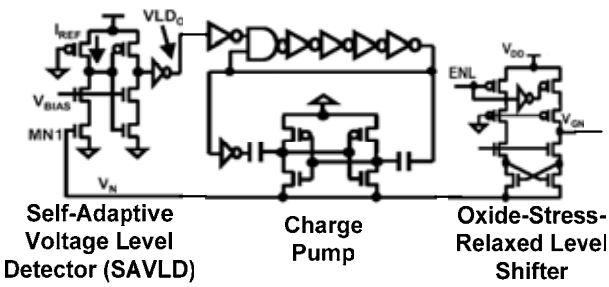


Figure 4. V_N (negative voltage) generator for NMOS cut-off switch with self-adaptive voltage level detector, charge pump, and oxide-stress-relaxed level shifter

In the proposed Optimal Zigzag (OZ) approach, a target circuit is synthesized with dual-V_{TH} scheme first. Then, an optimal input vector for IPF is identified so that the number of off-off gates on critical path is maximized. A random-based Monte Carlo search is effective in this step [3]. Finally, ZSCCMOS power switches are inserted for low-V_{TH} cells selectively in gate-by-gate manner instead of inserting one big power switch for a block of gates. This gate-by-gate insertion of power switch eliminates difficulty in timing closure. After the input phase forcing to achieve optimum off-off stack gates, the exact output values of low-V_{TH} cells in a standby mode are determined uniquely. Thus, either header or footer is determined uniquely according to the output value of each gate.

The proposed scheme has less overhead than the SMT because the keeper circuit can be excluded. The delay overhead is also smaller than the SMT because more than 30% of cells in a critical path can be off-off stacked in a standby. The off-off stack does not need any power switch and thus the delay overhead is eventually zero over the dual-V_{TH} design. There is overhead for modifying the first gates to input phase forcing gate but the total delay overhead is less than 2% on average which is shown in Table I.

III. Implementation and Results

The proposed OZ scheme is compared with many other candidates for low leakage design by using environments based on C/C++/STL and Perl on Unix machine. Industry-standard commercial CAD tools are employed for logic synthesis, spice simulation, functional simulation and layout. Ten ISCAS benchmark circuits were used for the experiments. The target technology is 65nanometer. Six different set of technology library (low-V_{TH} , high-V_{TH}, IPF, MT, SMT, OZ) are generated. Each library set has 173 cells. Each benchmark circuit has the following structures; c432 (152/18, total number of cells/depth), c499 (457/14), c880 (351/16), c1355 (586/26), c1908 (419/30), c2670 (557/20), c3540 (845/21), c5315 (1363/22), c6288 (3845/29), and c7552 (2009/27).

First, the benchmark circuits were synthesized with dual-V_{TH}. The percentage of low-V_{TH} cells is determined by area and delay constraints. Fig.5 shows the percentage of the low-V_{TH} cells for each benchmark circuit. 23% of the cells on average (min: 9%, max: 48%) were remained at low V_{TH} as shown in Fig.5. Then, the IPF optimization was conducted on critical paths and 36% of the cells (min: 28%, max: 47%) in a critical path were off-off stacked as shown in Fig.6.

Fig.7 shows results on random-based input vector generation and its convergence characteristics. Amazingly several hundreds of trial vectors are sufficient to achieve good optimization which was first pointed out in [3]. This makes the optimization time within several minutes and makes the approach feasible. After the IPF step, the power switch is inserted to each of the remaining low-V_{TH} cell that is not off-off stacked even after the careful choice of the IPF vector.

As to the size of the power switch, Spice simulation is performed to determine the good size as shown in Fig.8 in terms of delay and leakage current. The over-drive voltage is either 0V or 0.1V. Even when the over-drive voltage is zero when there is no need to generate any negative voltage, one order of magnitude improvement of leakage is observed because of the off-off stacking effect [8]. As a result of Fig.8, the power switch width is selected twice the size of the normal transistor width in an original gate.

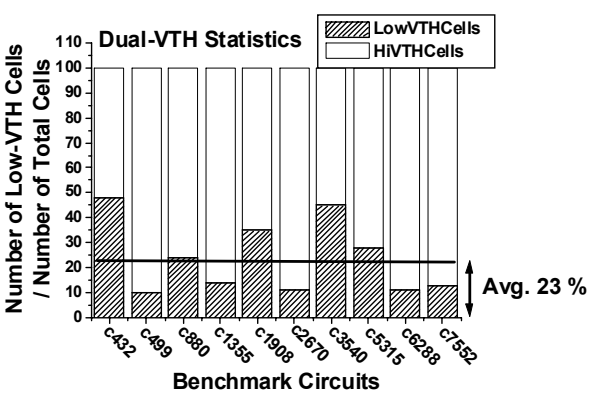


Figure 5. Dual-V_{TH} statistics

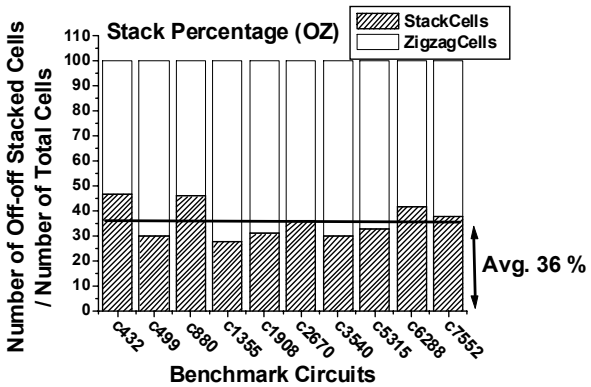


Figure 6. IPF results on critical path

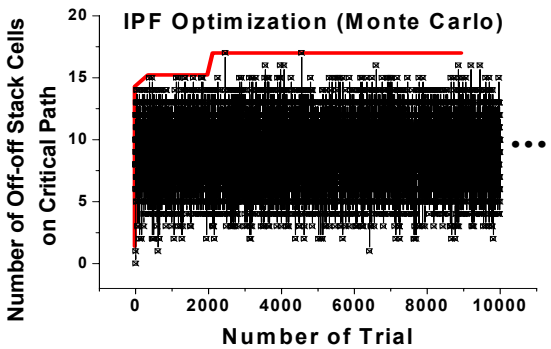


Figure 7. Optimization for IPF

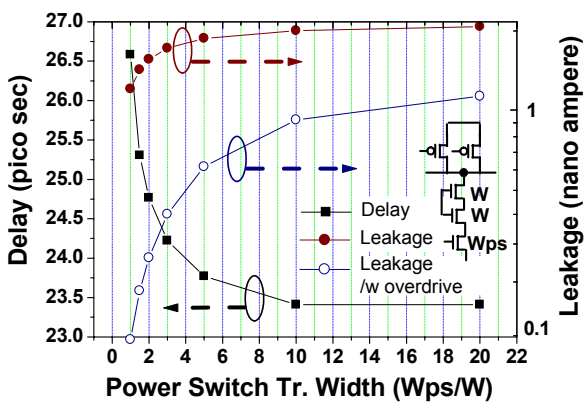


Figure 8. 2NAND vs. zigzag switch width

Figures 9, 10, 11 show the comparison results for ten benchmark circuits in terms of leakage, delay, and area, respectively. Note that ORG denotes all low-V_{TH} design, HVT, all high-V_{TH}, DUAL, dual-V_{TH}, MT, original MTCMOS, IPF, input phase forcing only, SMT, selective MTCMOS, and OZ denotes the proposed optimal zigzag scheme. The numerical values are tabulated in Table I. What is to be noted is that by using the OZ scheme, 70% leakage, 9.4% delay and 37% area reduction on average are achieved over selective MTCMOS which is the most widely adopted power-gating approach in the industry. Fig.12 shows an example of layout of each cell and Fig.13 shows a whole circuit layout and area comparison example for a 16x16 multiplier (c6288) case.

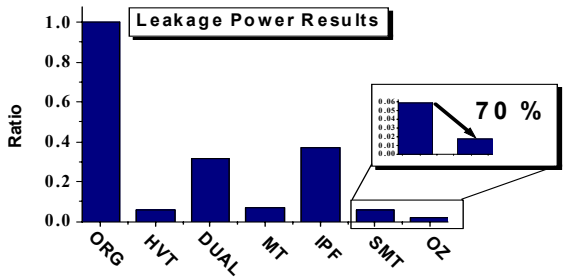


Figure 9. Leakage comparison Results

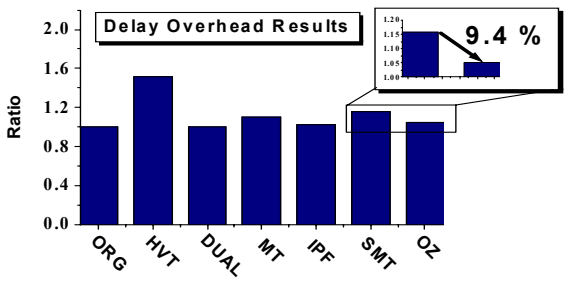


Figure 10. Delay comparison Results

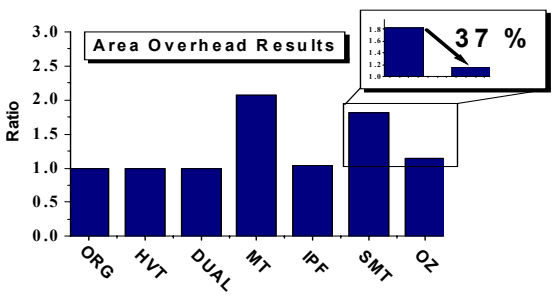


Figure 11. Area comparison results

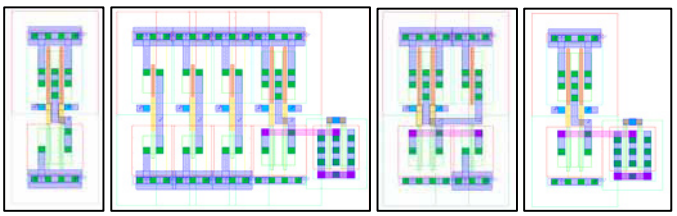
TABLE I. Average results for ten benchmark circuits

- 10 benchmark: c432, c499, c880, c1355, c1908, c2670, c3540, c5315, c6288, c7552
- Tech.: 65nm (Vdd : 0.9V, diff-VTH: 0.1V, sleep tr. size: 2W, overdrive : 0.1V)
- % of Low VTH cells / total cells: 23.4 % on average (min.: 9.19% / max.: 48.01%)
- % of IPF on critical path : 36.4 % on average (min.: 28.01% / max.: 47.05%)
- % of IPF over Low Vth cells : 20.1 % on average (min.: 11.9% / max.: 30.12%)

Power Gating Scheme	Description	Normalized by ORG		
		Power(leak)	Delay	Area
ORG	All Low-VTH (Original Circuit)	1.0	1.0	1.0
HVT	All High-VTH (Diff-Vt: 0.1 volt)	0.056	1.510	1.0
DUAL	Dual VTH (% of Low VTH: 23.4%)	0.315	1.001	1.0
MT	Gate-Level MTCMOS (2W-sleep tr. sizing)	0.069	1.096	2.072
IPF	Input Phase Forcing Only (Off-off Stacking only)	0.372	1.018	1.034
SMT	Selective MTCMOS (DUAL + MT)	0.058	1.160	1.817
OZ (No overdrive)	Proposed Optimal Zigzag (DUAL + IPF + ZZ)	0.087	1.049	1.138
OZ (Overdrive)	Proposed Optimal Zigzag (Overdrive Voltage: 0.1V)	0.017	1.051	1.152

IV. Conclusions

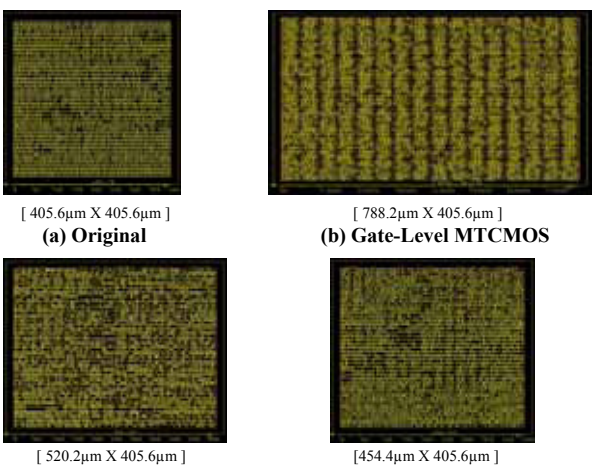
In this paper, an effective yet easy-to-implement power-gating scheme called OZ scheme to achieve two orders of magnitude lower standby leakage is proposed. The OZ scheme achieves 70% leakage reduction, 9.4% delay reduction and 37% area reduction on average are achieved over selective MTCMOS which is the most widely used power-gating approach to date.



[5.2μm X 13μm] [15.1μm X 13μm] [7.6μm X 13μm] [7.9μm X 13μm]

(a) NAND2 (b) NAND2-MT(latch) (c)NAND2-IPF (d)NAND2-Zigzag

Figure 12. Each cell design



[405.6μm X 405.6μm] [788.2μm X 405.6μm]

(a) Original (b) Gate-Level MTCMOS

[520.2μm X 405.6μm] [454.4μm X 405.6μm]

(c) Selective MTCMOS (d) Proposed Optimal Zigzag

Figure 13. Layout comparison (c6288, 0.18μm Tech.)

References

[1] L.Wei, Z.Chen, M.Johnson, and K.Roy, "Design and optimization of low voltage high performance dual threshold CMOS circuits," DAC'98, pp.489-494, Jun.1998.

[2] S.Mutoh, T.Douseki, Y.Matsuya, T.Aoki, S.Shigematsu, and J.Yamada, "1-V power supply high-speed digital circuit technology with multi threshold-voltage CMOS," IEEE J. Solid-State Circuits, vol. 30, pp. 847-854, Aug. 1995.

[3] J.P.Harter and F.Najm, "A gate-level leakage power reduction method for ultra-low-power CMOS circuits," CICC'1997, pp.475-478, Oct. 1997.

[4] K.Usami, N.Kawabe, and M.Koizumi, "Automated selective multi-threshold design for ultra-low standby applications," ISLPED'02, pp.202-206, Aug. 2002.

[5] K.Min, H.Kawaguchi, and T.Sakurai, "Zigzag super cut-off CMOS (ZSCCMOS) block activation with self-adaptive voltage level controller: An alternative to clock-gating scheme in leakage dominant era," ISSCC'2003, pp.400-401, Feb. 2003.

[6] J.Kao, S.Naredra, and A. Chandrakasan, "MTCMOS hierarchical sizing based on mutual exclusive discharge patterns," DAC'98, pp. 495-500, Jun. 1998.

[7] G.Uvieghara, et. al., "A highly-integrated 3G CDMA2000 1X cellular baseband chip with GSM/AMPS/GPS/Bluetooth/Multimedia Capabilities and ZIF RF support," ISSCC'2004, pp.422-423, Feb. 2004.

[8] T.Miyazaki, T.Q.Cahn, H.Kawaguchi, and T.Sakurai, "Observation of one-fifth-of-a-clock wake-up time of power-gated circuit," CICC'2004, pp.87-90, Oct. 2004.