

Energy Efficiency of Power-Gating in Low-Power Clocked Storage Elements

Christophe Giacomotto¹, Mandeep Singh¹, Milena Vratonjic¹,
Vojin G. Oklobdzija¹

¹ Advanced Computer systems Engineering Laboratory,
University of California, Davis,
One Shields Ave, Davis, CA 95616, USA,
 {christophe, mandeep, milena, vojina}@acsel-lab.com

Abstract. An analysis of the efficiency of power-gating for clock storage elements (CSE) is presented. Two CSE topologies: the TGMS (Transmission Gate Master Slave) and the WPMS (Write Port Master Slave) are examined along with their respective circuits with sleep transistors. In this work, we study the benefits of adding sleep transistors coupled with regular clock-gating during inactive mode. We examine the energy savings for standard clock gated CSEs versus their power gated counterparts. This is done by studying how the leakage energy saved with power gating offsets the energy consumed by the extra transistors added to support it. It is not always beneficial to add sleep transistors when deciding between power-gating or just using clock-gating. We also study how the results and tradeoff change with voltage scaling.

Keywords: Clocked storage elements, energy optimization, flip-flops, VLSI, power consumption, circuit tuning, circuit optimization, circuit analysis, power gating, clock gating, MTCMOS, voltage scaling.

1 Introduction

The traditional approach of scaling supply voltage for low-power applications becomes less energy efficient as technology continues to scale down. This is because leakage power is becoming more prominent and in low-power applications with a high standby-to-active ratio, it may become the dominant part of the total power consumption. Several methods have been developed to reduce the active energy of flip-flops [1][2][15][16]. However, in our analysis we will focus on leakage control methodologies on Master-Slave latch structures since these topologies are the most energy efficient in low power designs [14]. Circuit-level approaches to leakage power reduction include multi-threshold CMOS (MTCMOS) [6][7] and power-gating [8][9][10]. MTCMOS methodology utilizes high- V_T devices to decouple the low- V_T logic block from the voltage supplies during standby or sleep periods. Introduction of sleep transistors to aid leakage reduction is a common practice that presents some

trade-offs in terms of performance and active energy. Sleep transistors which decrease leakage through transistor stacking have shown to offer considerable reduction in leakage current. During sleep mode, the state of all circuits disconnected from the voltage supplies is lost. However, the state of all CSEs in a design needs to be preserved during sleep mode. Prior techniques for state retention in clocked storage elements during sleep mode were based on shadow copies, traditionally implemented with a balloon latch built of high- V_T devices, with a provided path to move data to and from the balloon latch [9]. Another technique is to use leakage feedback flip-flop (LFBFF) [6][10], which retains its state during sleep mode by conditionally maintaining an active path to voltage supplies based on feedback from the output.

In this work we examine common power saving techniques and the benefits of power-gating in the case of low power CSEs. Power-gating helps in effective leakage reduction but the overhead associated within the CSE can make the standard CSEs preferable. There exists a tradeoff between these techniques which we study in detail. The analysis methodology is presented in the context of clocked storage elements. In the following section the basic idea behind each of the techniques is presented followed by the circuits we examine in section 3. Section 4 details the influence of both clock- and power-gating which defines the energy efficient regions of operation, shown in section 5, for each of the analyzed methods.

2 Clock and Power Gating Methodologies

Fig. 1. depicts the setup arrangement used in our analysis. Clock-gating is controlled with “Enable” signal. There are two cases: a) the CSEs are modified such that power-gating can be applied. When the power-gating control signal “sleep” is turned ON their leakage current is expected to go down; b) the CSEs are traditional and not modified for power-gating leakage control; their leakage cannot be reduced during inactive time but they are more efficient in active mode.

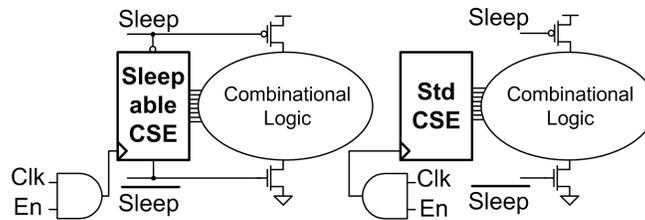


Fig. 1. Analysis setup: a) Power-gated CSE, b) Standard CSE

2.1 Clock Gating

In low power designs the clock distribution is typically gated at some level of the tree to reduce the active power consumed by the clock. In both cases shown in Fig. 1. the CSEs are clock-gated. We assume that when the “sleep” signal goes ON the clock is

also disabled. Because a large part of the clock power consumption comes from the capacitive load of the CSEs connected to the clock tree, in low power designs the clock-gating is done very close to the CSEs to allow a maximum control over the clock active power. Clock-gating can also help to reduce leakage power, which has an exponential dependence on the temperature [11]. Since dynamic power is reduced, the local temperature will decrease, thus resulting in lower leakage power.

2.1 Power-Gating

Leakage power increases with technology and voltage scaling to unacceptable levels [1][5]. To reduce leakage power, power-gating technique was proposed for clocked storage elements [6][9][10]. The sleep signal controls header and footer devices which are added between actual voltage supplies. These devices are turned-off in the sleep mode to cut-off the leakage paths in the combinational logic blocks and CSEs. There is an associated delay and power overhead in the CSE which is now modified for sleep mode operation. The penalties of modifications differ in different CSE topologies, which we investigate in this paper for two traditional Master-Slave latch structures. The main goal of our analysis is to explore the benefits and effective regions of power-gating versus solely clock-gating techniques for low-power CSEs.

3 Selection of Latches for Methodology Comparison

To demonstrate this analysis and selection methodology, we limited the choice of latches to two frequently used low power CSE topologies: the Transmission Gate Master-Slave latch (TGMS, Fig. 2a, [12]) and the Write Port Master-Slave latch (WPMS, Fig. 3a, [13]). The TGMS is a conventional master-slave latch topology in which both master and slave latches are implemented with transmission gates [1]. It is used in the Power PC 603 low-power processor [12], and it is generally considered as a low power storage element topology [14]. The WPMS works also like a master-slave structure. However, the implementation of each latch is inspired by a standard SRAM 6T cell. Each side of the keeper is controlled by a single nMOS which is driven by the clock or its complement if the latch is slave or master, respectively. When the clock opens both nMOS transistors in a wordline manner, the keeper is push-pulled from each side to change its state. The advantage of this structure is the removal of the pMOS from the pass-gates, which decreases the clock load and the parasitic capacitance on the data-path. However, since the pMOS is missing, the keepers cannot be conditional on the pull-up in order to bring the nodes *ma* and *sl* from $V_{DD}-V_T$ to V_{DD} when logic high is needed. Both M-S latch structures can be modified to support power-gating while retaining their internal states during sleep mode. In the case of the TGMS, the state retention can be done by using the leakage feedback methodology [6][7]. The resulting CSE is the TGMS_{SLEEP} (Fig. 2b) or canary flip-flop [10]. In the case of the WPMS, the state retention is intrinsic: since the state elements are not part of the data path in both the master and slave stages, the keepers can be implemented in high- V_T devices only. Consequently, as shown in Fig. 3b (WPMS_{SLEEP}), only the low- V_T inverters (I1, I2 and I3) need to be power-gated.

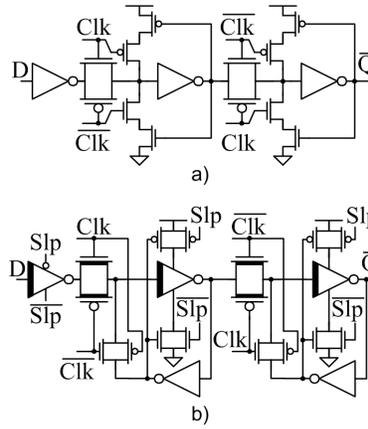


Fig. 2. a) Transmission Gate Master-Slave latch (TGMS), b) Power-gated TGMS latch [6] (TGMS_{SLEEP}) The low- V_T devices are highlighted

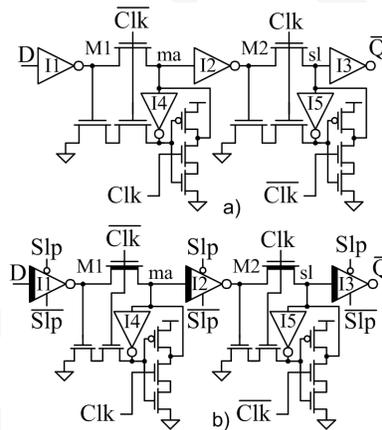


Fig. 3. a) Write Port Master-Slave latch [13] (WPMS), b) Power-gated WPMS latch (WPMS_{SLEEP}) The low- V_T devices are highlighted

4 Leakage vs. Active Power Contribution and Scaling

In order to illustrate the tradeoff between clock-gating and power-gating through the addition of sleep transistors we examine the contribution of both active and leakage energy components in the final analysis which defines the region of the most energy

efficient operation. The results presented in this section and section 5 were obtained by using HSPICE and 65nm CMOS technology models [17]. The setup, similar to the one used by [3][14], was used for the simulation of the individual CSEs during active mode.

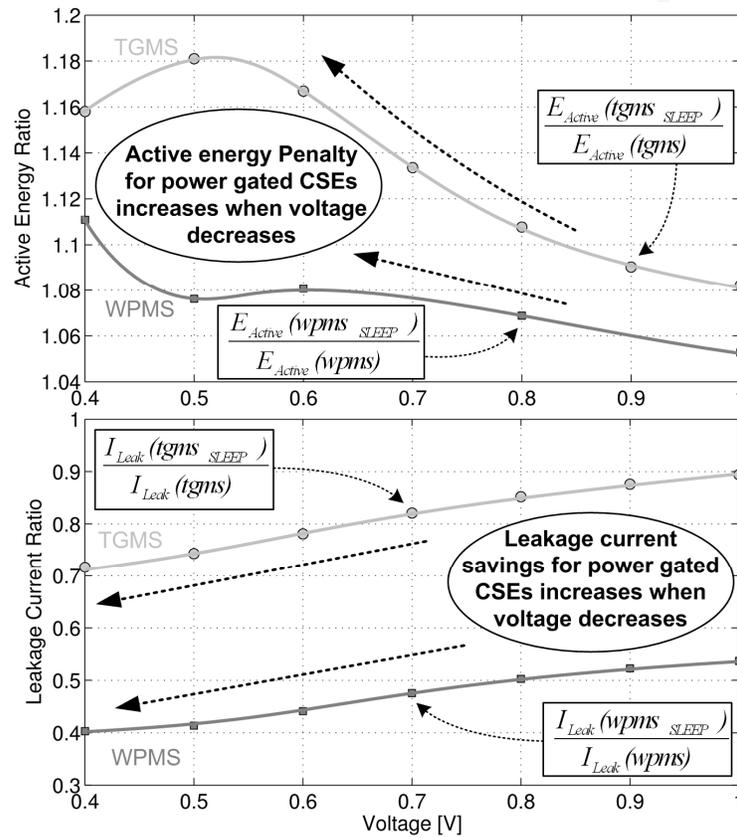


Fig. 4. a) Comparison of active energies between WPMS & WPMS_{SLEEP} and TGMS & TGMS_{SLEEP}, b) Comparison of leakage current between WPMS & WPMS_{SLEEP} and TGMS & TGMS_{SLEEP}

4.1 Leakage Energy

Fig. 4.b shows the savings in leakage current for both TGMS and WPMS CSEs with their respective power-gated circuits. Each curve is a ratio of the leakage current of the power-gated circuit over the standard clock-gated circuit. Lowering the voltage increases the savings in both latches linearly. The difference in the slope offset is due to the structure of the circuits.

In WPMS, the keeper for the state is separated from the main D to Q path, unlike in the TGMS. This structure allows power-gating to work better with the critical path low- V_T devices and helps to reduce the leakage considerably compared to the standard WPMS. The TGMS latch we examine does not have a decoupled keeper. If this latch is modified for power-gating the leakage on the critical path is not controlled as effectively due to the prerequisite of keeping the state of the latch. Extra transistors are required to achieve this and these in turn do not allow the low- V_T path to be completely gated. This results in less efficient savings in leakage current.

4.2 Active Energy

Although the leakage energy can contribute significantly to the total energy consumption, it is at least an order of magnitude lower than the active energy of the circuit until the supply voltage drops close to threshold voltage levels. Fig.4.a shows the ratio of active energies for each of the latches. This figure shows how the active energies of the power-gated circuits decrease at a slower pace than their traditional non-power-gated counterparts as we scale down voltage (both CSEs are clocked and data activity is 25%). This results in an increasing ratio of the active energies. Due to the circuit overhead, the power-gated CSEs consume more active energy than the standard CSEs. This overhead consists of additional transistors stacked in the data-path which reduces data-path efficiency at low voltages. However, this effect only offsets the ratio of active energies. The voltage dependency, i.e. the increasing ratio as voltage decreases, is due to the leakage because the data activity is 25% .

The following section, section 5, describes how the selection for the proper energy reduction methodology can be made. It is also shown how the dependencies described in section 4 influences the regions of energy efficient operation for each of the methods.

5 Energy Efficient Operation Regions and Selection Process

Fig. 5. illustrates the energy comparison between standard clock-gated WPMS latch with and without power-gating circuitry. The curve represents the break-even point for using standard WPMS versus the power gated WPMS at each voltage for minimum energy. The “percentage ON time” represents the amount of the time the CSEs are functional, by enabling clock in both cases and disabling sleep mode in the power-gated design. Therefore, at 0.5V if the WPMS were to be placed in a block which was expected to be active for 27% of the time, it is equally effective to use either clock-gating only (using the standard WPMS) or power-gating (using the WPMS_{SLEEP}) in order to achieve minimum energy.

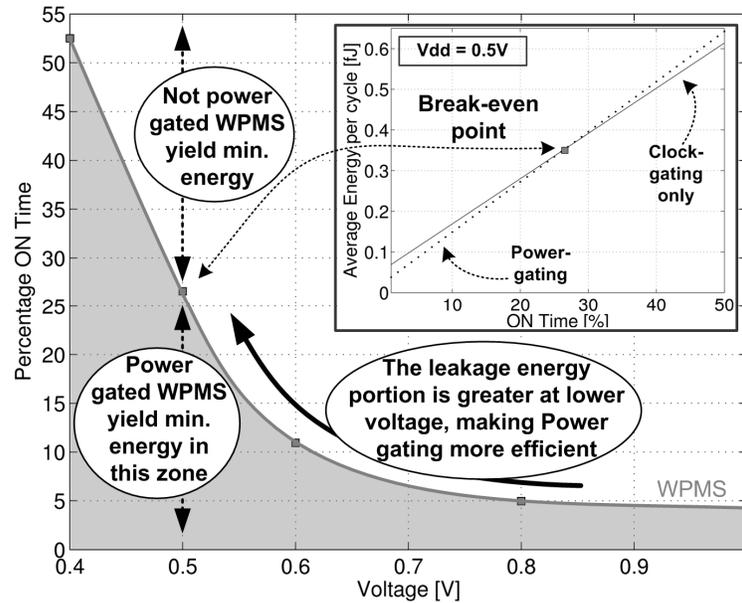


Fig. 5. Energy efficient zones of operation for either WPMS or $WPMS_{SLEEP}$ as a function of voltage scaling and ON time (amount of time the circuits are turned ON, in %)

If the circuit is active for more than 27% of the time, then clock-gating is preferable since the overall energy is lower than introducing the overhead associated with sleep mode transistors. Below 27% ON time lies the region of operation where it is more beneficial to use sleep mode transistors since clock-gating does not cut down leakage current enough to save energy during sleep mode. Thus, the region above the curve represents the operation zone where using sleep transistors is less effective than using the standard WPMS in reducing energy; the opposite holds true below the curve.

As the voltage is scaled lower and leakage becomes more significant, the relationship shown in Fig. 5. scales accordingly in favor of sleep transistors.

The inset plot in Fig. 5. shows the energy curves for the WPMS clock-gated and the WPMS power-gated at 0.5V supply voltage. These two curves are crossing over at the 27% ON time operation point, thus illustrating that this is the breakeven point at 0.5V for these two designs. It shows how the designs behave at 0.5V for a range of active operation times (ON times) confirming the breakeven curve in Fig. 5.

Using the same methodology as in Fig. 5, we analyzed the TGMS in the same context. The results are shown in Fig. 6. The operation zone for the $TGMS_{SLEEP}$ is relatively smaller than that of the $WPMS_{SLEEP}$ in Fig. 5. For example, at 0.5V the breakeven point for minimum energy is at 27% ON time between WPMS & $WPMS_{SLEEP}$ while it is at 3% for TGMS & $TGMS_{SLEEP}$. This means it is much less favorable to use the power-gating method for TGMS than for WPMS. This happens for two reasons:

- 1) The feedback leakage method [6][7] implemented in $TGMS_{SLEEP}$ creates additional leakage paths from the low- V_T devices used in its critical path which leads to less leakage reduction during sleep mode
- 2) The WPMS relies on the drive strength of the first stage to drive the master node low against the drive of the keeper pMOS during active operation. Since the keeper is made weaker with high- V_T devices, $WPMS_{SLEEP}$ becomes more efficient in the case of power gating.

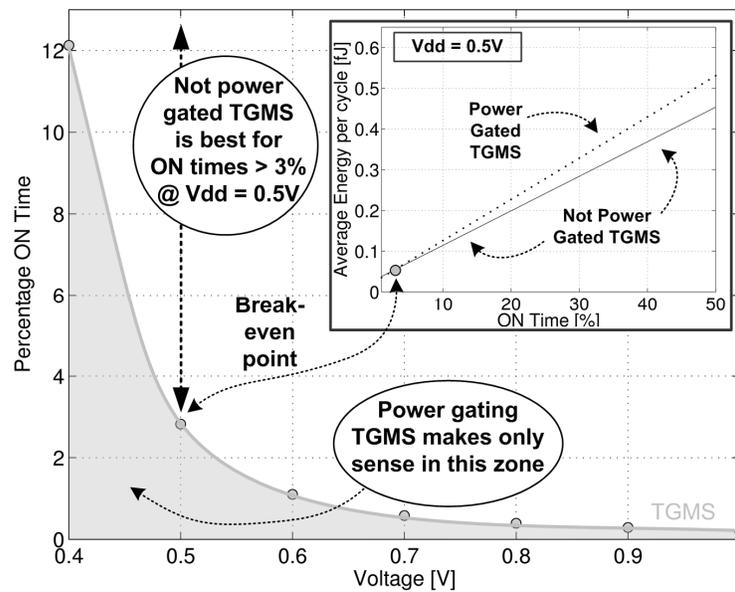


Fig. 6. Energy efficient zones of operation for either TGMS or $TGMS_{SLEEP}$ as a function of voltage scaling and ON time

6 Conclusion

This work presents a methodology for the analysis of energy efficiency of power-gated CSEs versus traditional CSEs. We show that depending on the structure, the operating voltage and the percentage of run time expected from circuits, power-gated CSEs may or may not be the optimum choice. Furthermore, we provided a novel and practical way to visually analyze the zones of energy efficient operation for CSEs, with and without power-gating for low power logic design.

References

1. V. G. Oklobdzija, V. M. Stojanovic, D. M. Markovic, N. M. Nedovic: Digital System Clocking: High-Performance and Low-Power Aspects, 1st Ed. Wiley IEEE-press (2003)
2. V. G. Oklobdzija, Ram K. Krishnamurthy: High-Performance Energy-Efficient Microprocessor Design (Series on Integrated Circuits and Systems), 1st Ed. Springer (2006)
3. V. Stojanovic and V. Oklobdzija: Comparative Analysis of Master-Slave Latches and Flip-Flops for High-Performance and Low-Power Systems, IEEE JSSC, Vol. 34, pp. 536—548, April (1999)
4. S. Borkar: Design Challenges of Technology Scaling, IEEE Micro, Vol. 19, No. 4, pp. 23—29 (1999)
5. K. Bernstein, C. T. Chuang, R. Joshi, R. Puri: Design and CAD Challenges in Sub-90nm CMOS Technologies, ICCAD, pp. 129-136 (2003)
6. J. Kao, A. P. Chandrakasan: MTCMOS Sequential Circuits. In: IEEE European Solid-State Circuits Conference (ESSCIRC), pp. 317—320 (2001)
7. B. H. Calhoun, F. A. Honore, A. P. Chandrakasan: A Leakage Reduction Methodology for Distributed MTCMOS, IEEE JSSC, Vol. 39, No. 5, pp. 818—826, May (2004)
8. M. Powell, S. Yang, B. Falsafi, K. Roy, T. Vijaykumar: Gated-Vdd: A Circuit Technique to Reduce Leakage in Deep-Submicron Cache Memories. In: IEEE International Symposium on Low-Power Electronics and Design (ISLPED), pp. 90—95 (2000)
9. S. Shigematsu, S. Mutoh, Y. Matsuya, Y. Tanabe, J. Yamada: A 1-V High-Speed MTCMOS Circuit Scheme for Power-Down Application Circuits, IEEE JSSC, Vol. 32, No. 6, pp. 861—869, June (1997)
10. B. H. Calhoun, A. P. Chandrakasan: Standby Power Reduction Using Dynamic Voltage Scaling and Canary Flip-Flop Structures, IEEE JSSC, Vol. 39, No. 9, pp. 1504—1511, Sept (2004)
11. S. Heo, K. Barr, K. Asanovic: Reducing Power Density through Activity Migration, ISLPED (2003)
12. G. Gerosa, et. al.: A 2.2W, 80MHz Superscalar RISC Microprocessor, IEEE JSSC, Vol. 29, No. 12, pp. 1440—1454, Dec (1994)
13. D. Markovic and J. Tschanz: Transmission-Gate Based Flip-Flop, US Patent 6,642, 765, Issued: Nov (2003)
14. C. Giacomotto, N. Nedovic, V. G. Oklobdzija: The Effect of the System Specification on the Optimal Selection of Clocked Storage Elements, IEEE JSSC, Vol. 42, No. 6, pp. 1392—1404, June (2007)
15. N. Nedovic, M. Aleksic, V. G. Oklobdzija: Conditional Techniques for Low Power Consumption Flip-Flops. In: 8th IEEE International Conference on Electronics, Circuits and Systems (ICECS), pp. 803—803 (2001)
16. N. Nedovic, M. Aleksic, V. G. Oklobdzija: Conditional Pre-Charge Techniques for Power-Efficient Dual-Edge Clocking. In: IEEE International Symposium on Low-Power Electronics and Design (ISLPED), pp. 56—59 (2002)
17. U.C.B.D. Group. BSIM4.2.1 MOSFET Model: User's Manual. Dept. of EECS, Univ. of California, Berkeley, CA 94720, USA, 2002.