

# A 0.4V 0.5fJ/cycle TSPC Flip-Flop in 65nm LP CMOS with Retention Mode Controlled by Clock-Gating Cells

Ludovic Moreau, Rémi Dekimpe and David Bol  
ICTEAM Institute, Université catholique de Louvain, Louvain-La-Neuve, Belgium  
Email: {ludovic.moreau, remi.dekimpe, david.bol}@uclouvain.be

**Abstract**—In this paper, we propose a low-overhead solution to ensure contention-free data retention in clock-gated true single-phase-clock (TSPC) flip-flops (FF) at ultra-low voltage (ULV). It relies on a retention feedback loop added to the TSPC FF and controlled by the clock-gating module. When the clock is gated, the retention is enabled, which drives the FF in retention mode. This limits the energy overhead induced by the added feedback loop and makes the FF contention-free. Moreover, as several FFs typically share the same clock-gating module, the control signal generation overhead is also kept low.

The proposed 19T TSPC FF with retention mode was implemented as a standard cell in 65nm LP CMOS. The FF energy is 0.5fJ/cycle at 0.4V, from post-layout simulations and for a typical 25% activity factor, which is 62% reduction compared to the conventional 24T master-slave FF. Experimental validation of a prototyped Cortex-M0 testchip including the integration of the proposed FF into synthesis and place/route flow validates its robust operation at ULV.

**Keywords**—TSPC flip-flops, data retention, clock gating, ultra-low power, ultra-low voltage, CMOS

## I. INTRODUCTION

Flip-flops (FFs) play a major role in digital circuits. They are the basic components of sequential circuits and, as such, they are responsible for an important part of their area and power consumption. FF standard cells are thus attractive candidates for power/area optimization. As an example, let us consider a case study in 65nm LP CMOS for this work: an ARM Cortex-M0 CPU with small instruction and data memories, synthesized from standard cells (Fig. 1). In this circuit, FFs account for 40% of the area and 30% of the power.

The master-slave FF architecture is conventionally used for standard cells, especially at ULV thanks to its high robustness. TSPC FFs [1] are an interesting alternative to the conventional master-slave architecture. Indeed, they perform better in terms of area, leakage power and energy per cycle [2]. However, they suffer from a drawback due to their dynamic state storage: the transistor leakage results in a data loss when there is no clock edge for a given time period. This problem appears when TSPC FFs are clock gated and is more critical at ultra-low voltage (ULV) where circuits have to operate at lower frequency, which leaves more time between consecutive clock edges. This impedes the use of TSPC FFs for ultra-low power (ULP) circuits, which need to combine ULV operation with clock gating [3]. To overcome this issue, we propose a low-overhead solution to ensure data retention in clock-gated TSPC FFs at ULV. It is based on a feedback loop added to the TSPC FF for data retention, which is controlled by the clock-gating modules.

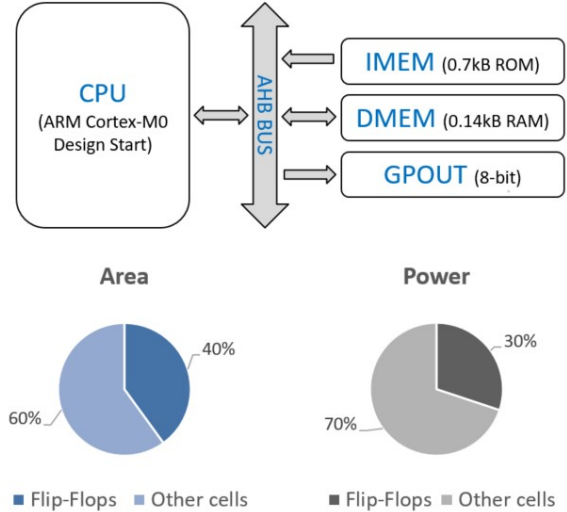


Fig. 1. Area and power breakdown of the case-study circuit in 65nm LP CMOS: a Cortex-M0 CPU with memories synthesized from standard cells.

## II. BACKGROUND AND MOTIVATION

In this section, we briefly recall TSPC operating principle before analyzing the data retention issue and how it is exacerbated at ULV.

### A. TSPC FF operation

The original TSPC [1] is composed of four inverting stages (Fig. 2). Hence, there are three internal nodes that we refer to as pre-charge ( $N1$ ), transfer ( $N2$ ) and retention ( $N3$ ). The first stage is responsible for the pre-charge of the data  $D$  to the node  $N1$  during the low phase of the clock, while the second stage PMOS always charges  $N2$  to a logic-1. At the rising edge of the clock, the second and third stages operate a data transfer from  $N1$  to the retention node  $N3$  by conditionally discharging  $N2$  to a logic-0, depending on the logic level of  $N1$ . This means that it stores the inverse of  $D$  in  $N3$ . Therefore, once  $N3$  is set, the fourth stage drives  $Q$  with the proper data by inverting  $N3$ .

### B. Data retention issue in clock-gated TSPC FFs

Automatic clock gating is a conventional low-power technique, widespread in digital circuits. It consists in holding low the clock of idle FFs in order to reduce their switching activity and therefore the overall dynamic power consumption. On top of good power saving results, the ease of adding clock gating in electronic design automation (EDA) tools makes this technique a must-do for ULP circuit design [3].

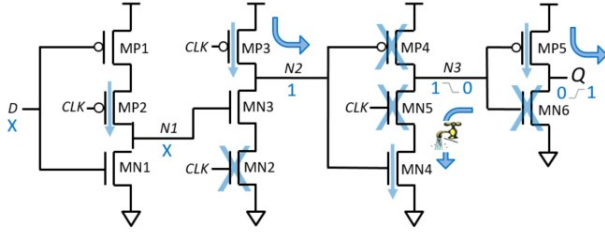


Fig. 2. Conventional TSPC FF. Additions in blue highlight the data retention issue that arises when clock-gated, due to leakage current in the third stage.

As mentioned previously, the original TSPC FF is likely to lose its data in absence of clock edges for a certain time. Let us briefly review why. When the clock is low,  $N2$  is forced to a logic-1 (Fig. 2). Thus, we do not care here about the first two stages and we can focus on the last two stages.  $N3$  stores the inverted data: in this case, a logic-1. Given that  $N2$  is high and  $CLK$  is low, the transistors  $MP4$  and  $MN5$  are OFF while  $MN4$  is ON. If this situation stays unchanged for too many cycles, leakage currents through  $MN5$  will slowly discharge  $N3$  to a logic-0, resulting in a data loss and a corrupted output.

### C. TSPC FF data retention at ULV

The example in Fig. 2 shows the loss of a logic-1 level in  $N3$  (hence a logic-0 data), assuming that the leakage current in  $MN5$  is stronger than  $MP4$ 's. Thus, the retention of a logic-0 level (hence a logic-1 data) is not an issue in this case. Obviously, there is a sizing tradeoff in the third stage that decides the critical logic level. Simulations in nominal conditions (1.2V, 500MHz) with the selected transistor sizes confirm this: the logic-0 data is lost after a time equivalent to 30k clock cycles while the logic-1 data suffers from a 10% decrease but remains stable even after more than 200k clock cycles. Let us mention that for all this work, the target clock cycle time ( $T_{cycle}$ ) is set as  $50 \times$  the fanout-of-4 (FO4) inverter delay [7] at the considered supply voltage.

There is no design point where this third stage would be perfectly balanced to offer logic-0 and logic-1 retention. Moreover, given the technology used (65nm CMOS), device variability cannot be ignored, especially at ULV. Indeed, nanoscale technologies suffer from higher variability due mainly to random dopant fluctuations and gate length variations [4]. These affect the transistor threshold voltage and make it a random variable, which induces a systematic PMOS/NMOS mismatch, even more critically at ULV. Therefore, when variability is considered, the TSPC retention of both logic levels is at risk, from one transistor to another. Fig. 3 shows Monte-Carlo SPICE simulations (1k runs) of the TSPC retention time of a logic-0 and a logic-1 in nominal conditions (1.2V) and at ULV (0.4V). Data loss occurs indeed in both cases, even if the logic-0 level remains more critical.

In addition, ULV implies a reduced clock frequency. As mentioned above,  $T_{cycle}$  is set to  $50 \times$  FO4 inverter delay, which gives a 1.14MHz target frequency at 0.4V. Consequently, despite the lower leakage currents compared

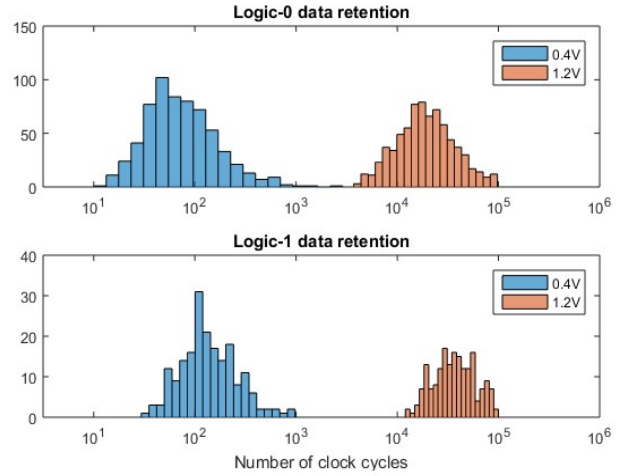


Fig. 3. TSPC FF data retention time distributions for a logic-0 and a logic-1 in the nominal case (1.2V, 500MHz) and at ULV (0.4V, 1.14MHz). Monte Carlo SPICE simulations (1k runs, TT 25°C corner).

to nominal conditions, the relative data loss time in terms of clock cycles is much lower. Fig. 3 highlights this difference, with a worst case going from 4k clock cycles (1.2V) to 10 (0.4V). To reinforce this measure, we used a statistical analysis methodology called gradient importance sampling (GIS) [6], with the following specifications: a target 99% chip yield with 10k FFs ( $5\sigma$  yield on a single FF). The analysis resulted in a worst-case retention time slightly above 4 clock cycles. As a side note, let us mention that with the GIS simulation methodology, only 14k simulation runs were needed compared to the theoretical 100M that Monte Carlo would require to obtain this result, corresponding to a  $7000 \times$  speedup. As this 4-cycle worst-case data retention certainly falls below typical clock-gating duration, it really motivates the need to add an explicit data retention mechanism in TSPC FFs for clock-gated ULV circuits.

## III. ENSURING DATA RETENTION IN CLOCK-GATED TSPC FLIPS-FLOPS

To ensure data retention in TSPC FFs, previous papers focused on adding always-on retention mechanisms in the FF [5]. However this generally requires many additional transistors which induce energy and area overhead. Hence, a different approach was chosen for this work, described in this section and compared to the master-slave architecture and then to state-of-the art static single-phase FFs.

### A. Proposed solution

To overcome the data retention issues, we propose a low-overhead solution based on the addition of a feedback loop at the retention node (Fig. 4a). This loop is composed of two inverters, making the TSPC a two-mode cell: normal mode (NM) when the clock is active and retention mode (RM) when the clock is gated. For that purpose, the second inverter is tristate so that we can disable the loop in NM and enable it in RM. Therefore, the loop is active only when it is needed and the benefit is twofold in NM: firstly, it limits the power overhead of the added transistors and secondly, it avoids contention at  $N3$  and guarantees a robust write operation.

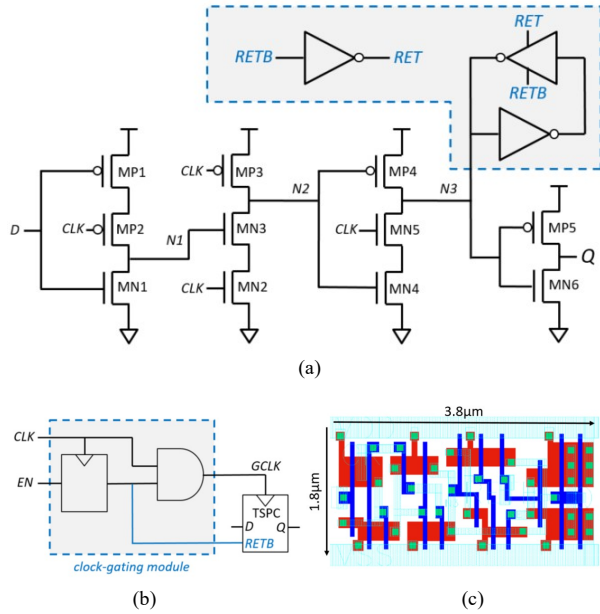


Fig. 4. Proposed solution to ensure data retention in clock-gated TSPC flip-flops at ULV: (a) a feedback loop is added at the retention node, controlled by (b) the clock-gating module. (c) Layout of the proposed TSPC with retention mode in 65nm LP CMOS.

Besides, as the mode switch corresponds to the gating of the clock, the loop control signal  $RETB$  can be directly issued by the clock-gating module, from the latched enable signal (Fig. 4b). This further limits the overhead of our solution as several FFs generally share the same clock. For example, in our case-study circuit, there is in average 1 clock-gating module for 11 gated FFs. Moreover, this solution is low-latency and ensures to enter RM before 4 clock cycles, which is the  $5\sigma$  worst case (Section II.C).

### B. Post-layout results

The proposed TSPC FF with retention mode was laid-out in 65nm LP RVT CMOS (Fig. 4c), following standard cells specifications. Table I compares it to the conventional master-slave FF, with post-layout simulation results at 0.4V. The simulation conditions [7] are:

- $T_{\text{cycle}}$  set as  $50 \times \text{FO4}$  inverter delay;
- a load  $C_L = C_{IVX4}$ , with  $C_{IVX4}$  being the input capacitance of an X4 inverter;
- a 25% activity factor ( $\alpha$ );
- a minimum data-to-output (D2Q) delay.

Results show a 62% reduction in energy per cycle, 26% reduction in leakage power and 5% reduction in area, with similar D2Q delay.

Fig. 5 provides post-layout simulations that illustrate the proposed TSPC operation in NM and RM at 0.4V and 1.14MHz, for 1k Monte-Carlo runs. The  $EN$  and  $CLK$  signals are the input of the clock-gating module (Fig. 4b) which provides the gated clock  $GCLK$  to the FF.  $RETB$  is the inverted  $RETB$  signal (Fig. 4a), hence the latched  $EN$ : the FF is in NM when  $RETB$  is low and in RM when  $RETB$  is high. These simulation waveforms show robust write operation in NM and data retention in RM.

Table II compares the proposed FF with two recent static single-phase FFs. Performances are normalized respectively with reported results of the master-slave FF. The results for this work arise from post-layout simulations while the one

TABLE I. PERFORMANCE COMPARISON AT 0.4V: 65nm LP/RVT CMOS POST-LAYOUT SIMULATION RESULTS

$T_{\text{cycle}}=50 \times \text{FO4}$ $C_L=C_{IVX4}$ $\alpha=25\%$	Standard Cell	This work
<b>Topology</b>	Master-Slave	TSPC with RM
<b>Number of transistors</b>	24	19
<b>Min. D2Q delay [ns]</b>	59.4	54
<b>C2Q delay [ns]</b>	45.7	34.6
<b>Setup time [ns]</b>	13.7	19.4
<b>Energy per cycle [fj]</b>	1.32	0.5
<b>Leakage power [nW]</b>	23.3	17.3
<b>Area [<math>\mu\text{m}^2</math>]</b>	7.20	6.84

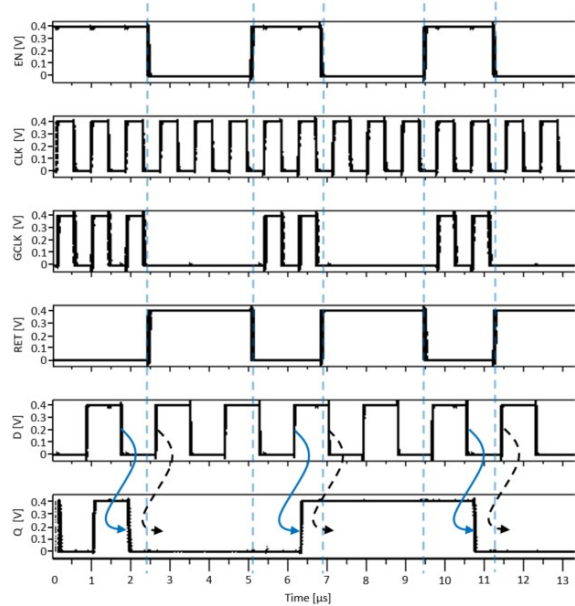


Fig. 5. Proposed TSPC FF operation in normal mode and retention mode. Post-layout Monte Carlo SPICE simulations with runs superimposed (1k runs at 0.4V, 1.14MHz, TT 25°C corner).

from the literature are testchip measures. While we implemented our FF in a testchip, our case-study circuit does not allow us to extract the performances of a single FF. However, we will demonstrate in Section IV that in terms of power, the measured testchips were mostly below the post-layout simulation result (Fig. 8). Power results in Table II can thus be seen a worst-case for the proposed FF. In the FF from [5], the greater transistor count and always-on loop implies higher energy per cycle and area than the proposed FF. On another hand, [2] have better normalized performances but at the cost of current contention on one node between strong pull-down NMOS and weak pull-up PMOS. This is a weakness for high- $\sigma$  statistical robustness. Moreover, it uses a mix of transistor lengths, which is a weakness in terms of design for manufacturability (DFM). Also, measurement results in [2] are reported for the FF with reset, for which the TSPC topology compares even better to the master-slave one. Finally, [8] shows very good energy per cycle at the expense of transistor count and area.

### IV. EXPERIMENTAL VALIDATION

To validate the proposed TSPC experimentally, we prototyped the case-study circuit from Fig. 1 in 65nm LP CMOS testchip with RVT transistors, packaged in QFN48.

TABLE II. PERFORMANCE COMPARISON AT 0.4V BETWEEN STATE-OF-THE-ART STATIC SINGLE-PHASE FFs (TESTCHIP MEASURES) AND THE PROPOSED TSPC WITH RETENTION MODE (POST-LAYOUT SIMULATION RESULTS)

$C_L=C_{IVX4}$ $\alpha=20\%$	[5] ISSCC'14	[2] TCAS-I'17	[8] JSSC'18	This work
<b>Topology</b>	S <sup>2</sup> CFF	Retentive TSPC	CSFF	TSPC with RM
<b>Retention</b>	Always-on	Always-on	Always-on	Controlled by CG module
<b>Contention-free</b>	Yes	No	Yes	Yes
<b>Technology</b>	PDSOI CMOS	FDSOI CMOS	Bulk CMOS	Bulk LP CMOS
<b>Node</b>	45nm	28nm	40nm	65nm
<b>Transistor type</b>	Single $V_T$ / min $L_g$	Single $V_T$ / multi $L_g$	Single $V_T$ / min $L_g$	Single $V_T$ / min $L_g$
<b>Transistors clock/total</b>	5/24	4/18	4/24	4/19
<b>Norm. C2Q delay*</b>	0.85	0.71**	0.73	0.76
<b>Norm. energy per cycle*</b>	0.63	0.43**	0.24	0.39

\* normalized to reported master-slave FF

\*\* results reported for the reset FFs and 100% activity factor

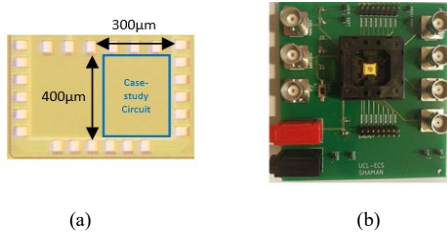


Fig. 6. (a) Microphotograph of the testchip (b) Experimental validation PCB.

As mentioned in Section I, this circuit consists of an ARM Cortex-M0 CPU with small instruction and data memories, synthesized from standard cells. Two versions of the proposed FF (with and without reset) were laid out and added to the standard cell library, which was characterized at 0.4V. Besides, custom automated steps were added post-synthesis to the EDA flow in order to tweak the clock-gating modules and wire them to the FFs. In total, 10 dies (Fig. 6a) were measured using the PCB in Fig. 6b. A simple counting software combined with the 8-bit GPOUT bus are used to monitor the behavior of the circuit under various supply voltage and frequency conditions. Fig. 7 reports the operating frequency range of all the dies at 0.4V. As the feedback loop is not always-on, the FFs have a low frequency bound. These measures are consistent with worst-case data retention from Fig. 3. Fig. 8a shows the measured power of the 10 testchips which are in range with post-layout simulation and mostly below it (3.3% in average), validating Table II results. Finally, Fig. 8b shows the minimum operating supply voltages, with the worst die at 0.33V, which demonstrates the robustness at ULV.

## V. CONCLUSION

In this paper, we propose a low-overhead technique to ensure contention-free data retention in clock-gated TSPC FFs for ULV circuits. This technique consists in the addition of a feedback loop at the retention node, conditionally enabled by the clock-gating module shared by several FFs. As the loop is active in retention mode only, there is no

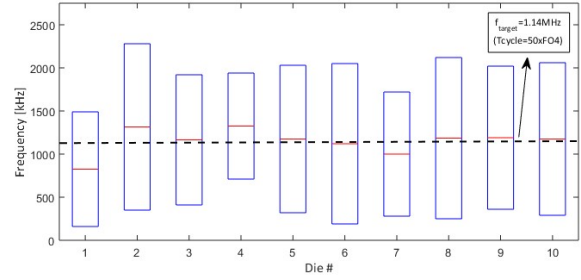


Fig. 7. Operating frequency range of the 10 measured testchips at 0.4V.

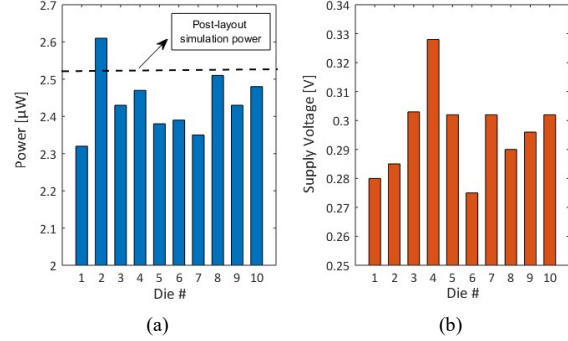


Fig. 8. Testchips measures: (a) power at  $V_{DD}=0.4V$ ,  $f=1.14MHz$  in regard to post-layout simulation and (b) minimum operating supply voltage.

contention. Post-layout simulations show a 62% energy-per-cycle gain compared to the master-slave FF, at 0.4V and for a typical 25% activity factor. Finally, the technique was validated with a 65nm LP CMOS testchip including a full Cortex-M0 synthesized from standard cells that demonstrated wide operating frequency ranges and a worst-case minimum supply voltage of 0.33V for 10 tested dies.

## ACKNOWLEDGMENT

The authors would like to thank Martin Lefebvre for the design of the experimental validation PCB.

## REFERENCES

- [1] J. Yuan and C. Svensson, "New single clock CMOS latches and flipflops with improved speed and power savings", IEEE Journal of Solid-State Circuits (JSSC), pp. 62-69, Jan. 1997.
- [2] F. Stas and D. Bol, "A 0.4-V 0.66-fJ/cycle retentive true-single-phase-clock 18T flip-flop in 28-nm fully-depleted SOI CMOS", IEEE Trans. on Circuits and Systems I: Regular Papers (TCAS-I), pp 935-945, Dec. 2017.
- [3] D. Bol et al, "SleepWalker: a 25MHz 0.4-V sub-mm<sup>2</sup> 7-µW/MHz microcontroller in 65-nm LP/GP CMOS for low-carbon wireless sensor nodes," IEEE J. of Solid-State Circuits, pp. 20-32, Jan. 2013.
- [4] D. Bol, R. Ambrose, D. Flandre and J-D. Legat, "Interests and limitations of technology scaling for subthreshold logic", IEEE Trans. on VLSI Systems (TVLSI), pp. 1508-1519, Oct. 2009.
- [5] Y. Kim et al, "A static contention-free single-phase-clocked 24T flip-flop in 45nm for low-power applications", IEEE International Solid-State Circuits Conference, pp. 466-467, 2014.
- [6] T. Haine, J. Segers, D. Flandre and D. Bol, "Gradient importance sampling: an efficient statistical extraction methodology of high-sigma SRAM dynamic characteristics", IEEE Design, Automation and Test in Europe (DATE), 2018.
- [7] M. Alioto, E. Consoli and G. Palumbo, "Analysis and comparison in the energy-delay-area domain of nanometer CMOS flip-flops: part II--results and figures of merit", IEEE Trans. on VLSI Systems (TVLSI), pp. 737-750, May 2011.
- [8] V. Loi Le, J. Li, A. Chang and T. Tae-Hyoung Kim, "A 0.4-V, 0.138-fJ/cycle single-phase-clocking redundant-transition-free 24T flip-flop with change-sensing scheme in 40-nm CMOS", IEEE J. of Solid-State Circuits, pp. 2806-2817, Aug. 2