

Area and Power-Delay Efficient State Retention Pulse-triggered Flip-flops with Scan and Reset Capabilities

Kaijian Shi
Synopsys (Professional Services)
Kaijian.Shi@synopsys.com

Abstract—This paper presents two area and power-delay efficient state retention pulsed flops with scan and reset capabilities for sub-90nm production low-power designs. The proposed flops also mitigate area overhead and integration complexity in SoC designs by implementing a single retention control signal and shared function/scan mode clock.

I. INTRODUCTION

The continually shrinking process feature size and increasing chip integration scale has resulted in power criticality in sub-90nm chips. The exponential increase of leakage with V_{th} reduction makes the power issue even more critical in battery-powered applications. Power-gating [1]-[9] is one of the most effective methods to reduce chip power in the sleep mode where inactive blocks are shut down by turning off power supplies to the blocks through switch cells. Critical states of the design can be shifted out and saved in memory before the blocks go to sleep and retrieved from the memory at block wakeup and before resuming operations. However, this state saving and restoring process could take as long as a few milli-seconds which could result in significantly performance degradation. To enable rapid resuming operations from the sleep mode, many power-gating designs implemented state retention registers to retain design states in the sleep mode. At wakeup, the retained states are quickly restored to allow the application to continue from the states in which the design went to sleep.

Various state retention registers have been developed [9]-[12]. They can be characterized into two types. In the first type of the retention registers, a bloom-style retention latch is added to a normal register. The latch is controlled by two retention signals; one saves the register state before going into the sleep mode and another restores the state after the register is waken up. The two signals are asserted only for a short period when going into sleep and waking up. They are kept low in the function and sleep modes to isolate the state retention logic part from the function mode register. This helps to reduce power overhead introduced by the retention logic to the register. However, the distributions of the two retention control signals through buffer trees introduce considerable power and area overhead in SoC designs where the buffer trees can be very large to distribute the signals to hundred thousands registers.

The second type of retention registers is built based on the master-slave style flip-flops. The slave latch is used for state

retention in the sleep mode. This is done by keeping the slave latch alive during the sleep mode. The clock must be held low during the sleep mode to prevent state in the slave latch from being corrupted through data path from the master latch which lost power in the sleep mode. Since the state has been in the slave latch during normal operation before going into the sleep mode, the state saving process becomes not necessary. Consequently, the state saving signal is removed; only state restoration control signal is needed in this type of registers. However, adding the clock holding logic in the clock paths introduces noticeable power and performance penalties. The clock-to-q delay becomes longer due to the introduced delay in the clock holding logic; the register power increase because the clock holding logic is in the high toggle rate clock path.

Both types of the retention registers are built on the master-slave type flip-flops which are not as efficient in area and power-delay product as the pulse clock triggered flip-flops [13]-[23] (pulsed flops for short). The pulsed flops are essentially latches operated on pulsed clocks. Therefore, they are smaller and faster than master-slave registers. It would be valuable to the low-power design industry, if pulsed flops with state retention, reset and scan capabilities were developed for production low-power designs. So far, only researchers at IBM Watson Research Center reported such retention and scan pulsed flops [24]. In their designs, the state retention technique was developed based on dual test clock controlled scan pulsed latches. The idea was to leverage the dual test clocks with the control of function clock to accomplish state save and restoration with an added state restoration control signal. The details of the designs are summarized in the next section. The shortcoming of such state-retention pulsed flops is the need of four global signals (three clocks and one RESTORE) to accomplish the state retention operation. Buffer tree distributions of four global signals to hundred thousand registers incur considerable overhead in area, power and routing resources.

This paper proposed two new state retention pulsed flop designs. The flops implement a single clock pin for function and test mode clocking. This simplifies clock distributions and reduces power in the clock trees. Moreover, the scan mode operation of the proposed designs is compactable to existing scan-based test designs. Therefore, it is feasible to leverage conventional DFT flow in a low-power designs that implement the proposed flops. The retention circuit in the

new state retention pulsed flops has been designed with minimum area overhead and controlled by a single retention control signal instead of dual “save” and “restore” signals in other retention flops. The reset circuit of the flops is also minimized consisting of only two transistors. As the result, the proposed flops are significantly area and power-delay efficient.

In the rest of the paper, the prior art of work reported by IBM researchers is outlined with comments on advantages and shortcomings. Then, the new state-retention pulsed flop with scan and reset capabilities is described in details with simulation results. Next, an area reduction version of the new state-retention pulsed flop is presented with circuit schematic. After that, the two proposed flops are compared with two production retention flops from two leading-edge low-power library providers.

II. PRIOR ART OF WORK

Various state-retention pulsed flops were reported by IBM [24]. These flops were designed based on dual test clock controlled scan pulsed latches. The principle of the retention technique in these designs is to leverage exiting two test clocks to control state saving and restoring when going into sleep and at wakeup respectively. The conceptual design of these flops is depicted in Fig. 1.

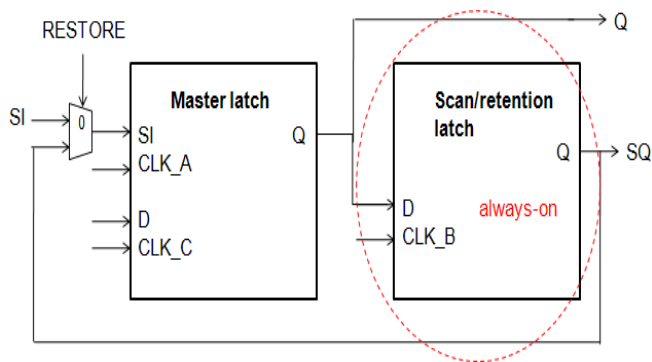


Fig. 1 Dual test clocks scan latch based retention pulsed flop

These state-retention pulsed flops are composed of a master latch and a scan latch. In normal operations, the two test clocks (CLK_A and CLK_B) are held low to block scan input to the master latch and D input to the scan latch. Consequently, the design becomes a conventional pulsed latch behaving like a flop clocked by pulsed clock CLK_C.

In scan mode, the function clock CLK_C and RESTORE are held low to select scan data input. CLK_A and CLK_B are non-overlapping scan clocks which clock master and slave latches respectively. Consequently, the flop behaves as a static master-slave flop in the scan mode.

In sleep (power-gating) mode, the state retention is operated in a three stage sequence.

When entering the sleep mode, CLK_A and CLK_C are

held low. CLK_B is pulsed to save the current state into the slave latch.

During the sleep mode, power supply to the design is shut off, except the slave latch which is connected to “always-on” power supply to sustain alive and retain the state. CLK_B is kept low to prevent state corruption from output of the master latch.

At the wakeup, the power supply to the flop is restored first, then RESTORE is asserted and CLK_A is pulsed to restore the state into the master latch through an added feedback path from the slave latch. CLK_B and CLK_C are kept low to prevent the data from being corrupted.

The main advantage of the designs is low area overhead. Only a mux and state-restoration feedback path controlled by RESTORE are added. However, the advantage only realizes in the dual test clock controlled scan pulsed flops to which the retention method becomes applicable. Consequently, these retention pulsed flops inherit the shortcomings of the scan pulsed flops. The master-slave latch pair is not area efficient compared with the pulsed latches. Moreover, the design requires four global control signals; three clocks and one retention control signal. Buffer tree distributions of these signals to all retention flops in a production design introduce considerable overhead in area, power and routing resources.

III. AREA AND POWER-DELAY EFFICIENT STATE RETENTION STATIC PULSED FLOP

To overcome the shortcomings of the retention pulsed flops in [24], a Low Power Testable Static Pulsed Flip Flop with Reset and Retention (LPTSPFFRR) has been developed for production low-power designs. It is worth noting that production low-power designs commonly require flops to have asynchronous reset capability to facilitate test, verifications and state control when the designs change operation modes. For this reason, a concise asynchronous reset circuit is designed in the proposed flops.

The circuit schematic of the design is shown in Fig. 2. The power supplies (VDD, Virtual VDD and GND) are provided and controlled by chip level power management unit and hence not shown in the circuit schematic.

The flop is clocked by a single pulsed clock in both function and scan modes. The switching between the function and scan mode clocks is controlled by chip clock management unit in a same way of conventional designs implementing static scan flops.

Two tri-state inverters are implemented at the inputs to select data from D pin in function mode and SI pin in test scanning mode by scan enable signal SE. The choice of the tri-state inverter instead of pass-gate is based on the reason that the pass-gate logic input flops are sensitive to input signal’s transitions resulting in large delay variations. This driving dependency and delay variation is costly to manage in production low-power designs. Only those push-to-the-limit CPU designs afford such effort. To avoid this issue, the

tri-state inverters are implemented in the designs.

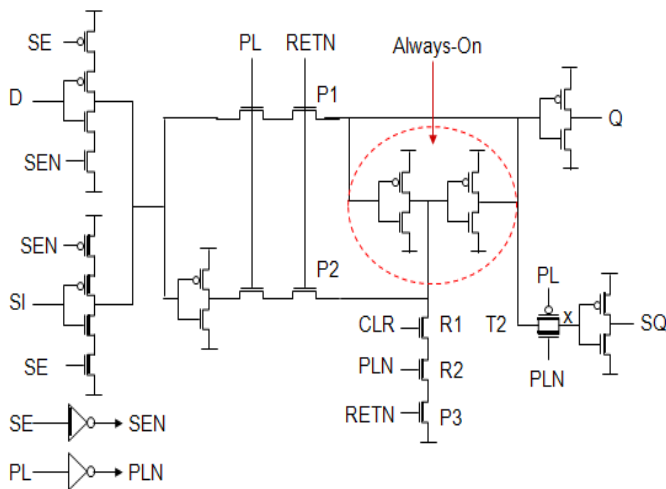


Fig. 2. Low Power State Retention Testable Static Pulse-triggered Flip-flop

The differential input latch structure is adopted in this design for its advantages of low clock power and permissible implementation of a small size latch inverter pair without significant impact on performance. Moreover, NMOS pass transistors instead of CMOS transmission gates can be used in the design to reduce area and power overhead on the clock paths. This becomes possible due to strong pull-down of the NMOS pass gates at one of the two latch differential inputs. The Q pin provides the shortest delay paths from clock pin (PL) and data pins (D and SI) for high performance.

The state retention circuit is designed using only three transistors (P1, P2 and P3) controlled by a single control signal “RETN”. In function and test modes, “RETN” is kept high which opens the three pass transistors it controls. Consequently, the pulsed flop functions as a testable and resettable flop.

When the flop is going into the sleep mode, “RETN” is pulled low (logic “0”) to close the three pass transistors (P1,P2, and P3). This isolates the latch (inverter pair) from the rest of the design. The latch is connected to permanent power supply to keep alive and retain state during the sleep mode when power supply to the flop is shut off.

At wakeup, the power supply to the flop is restored. Then “RETN” is pulled up to “1”. This opens the three pass transistors and the flop is back to functional operations with the retained state.

Transistor P3 is added in the reset path to protect the retained state from being corrupted by the asynchronous reset which could occur at wakeup before the flop resumes operation from the retained state.

The state save and restore sequence is illustrated by the control signal waveforms in Fig. 3.

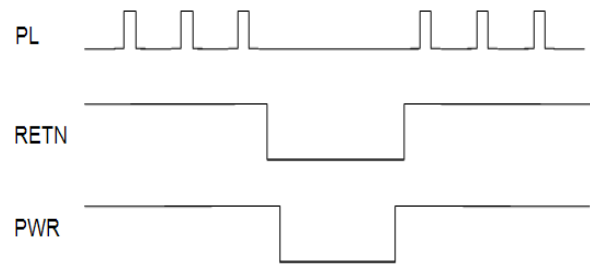


Fig. 3. State save and restore sequence

The reset circuit of the proposed flop has been designed with minimum overhead on power and area. Only two NMOS transistors (R1 and R2) are added to the design as shown in the circuit diagram in Fig. 2. The top pull-down NMOS transistor (R1) is controlled by the reset signal to force “0” state in the flop when the reset signal is asserted. The bottom pull-down NMOS (R2) is controlled by the pulse clock. It closes during the pulse window to prevent data contention between reset and input state. The pass transistor P3 in the reset pull down path is added to prevent state of the flop from being corrupted by the asynchronous reset at wakeup.

The scan shift function is implemented by adding a transmission gate (T2), controlled by the clock, to the SQ output path. T2 closes in latch opening window to prevent current state from passing through to the next flop in the scan chain. As the result, the pulsed flops can be back-to-back connected in the scan chain without hold time violations. It is worth noting that this hold-violation-free nature can also be leveraged in the function mode to map SQ pin, instead of Q pin, to those short logic paths between two pulsed flops. This avoids inserting hold-fix buffers in the paths and hence reduces the power and area overhead.

It is also worth noting that the design does not have the scan-out inverter-pair latch commonly implemented in the conventional static scan flops and the retention scan pulsed flops in [24]. The transmission gate (T2) in the proposed design is sufficient to latch the scan output, because T2 stays open most of the time during operations. It only closes for a very short period (pulse width) during which the junction capacitance of T2 and gate capacitance of the output inverter are sufficient to hold the signal on internal node which is marked X in Fig.2.

The proposed design has been optimized to minimize power-delay product in TSMC65G technology. The production constraint of minimum transistor size is considered during the transistor sizing.

It is worth noting that the functionality and IO behavior of the proposed pulsed flop are compatible to the conventional master-slave style retention scan flops. Therefore, it can be used to replace the conventional retention flops in existing designs to improve power-delay efficiency and chip area.

IV. SIMULATION RESULTS

The proposed design was evaluated by HSPICE simulations using TSMC65GPLUS (65nm) BSIM4 (v4.5) device models. A sleep transistor is added in the simulation deck to provide power-on and power-off supply to the flop in normal and sleep mode operations.

The simulation started in function mode ($SE=0$) for the first three cycles. Then, the RETN was asserted followed by CLR and PWRN (negative of PWR) to simulate state retention and power-gating behavior. Next, the PWR is restored, followed by CLR and de-asserting RETN to verify wakeup and state restoration of the flop. After that, the SE was asserted and data transitions were applied to SI inputs for three clock cycles to simulate scan shift function of the flop in scan mode. A CLR pulse was applied at the end of the simulation to check asynchronous reset behavior. The I/O waveforms of the simulations in function, sleep, wakeup and scan modes with the simulation deck above are captured and shown in Fig. 4.

Looking at the waveforms in the first three cycles, it is seen that flops' Q and SQ pins correctly follow the input signal transitions at D pin in function mode.

During the sleep mode, both Q and SQ are floating at about a half of VDD as the result of voltage distribution through the leakage paths of sleep transistor and the flop.

At the wakeup, the state of the flop was restored correctly as shown in Q and SQ waveforms when PWRN was de-asserted followed by CLR and RETN de-assertions.

In the scan mode, after the wakeup and SE assertion, the waveforms show that signals at Q and SQ correctly track SI signal instead of D inputs. Also, the output at SQ is delayed by a pulse width as designed to allow back-to-back flop connections in scan mode without hold violations.

As to asynchronous reset simulation, CLR was asserted twice in the simulations; one during the sleep mode to check effect on state retention and another close to the end of the simulation to verify asynchronous reset. It is seen from the waveforms that the state was not corrupted by the reset during the sleep mode nor during the period after the power is restored and before the retention control RETN was de-asserted. This is shown by Q and SQ states which were restored correctly at wakeup after RETN was lifted. At the end of the simulation, when CLR was asserted, the state was simultaneously reset to "0 as shown in the waveforms of Q and SQ pins.

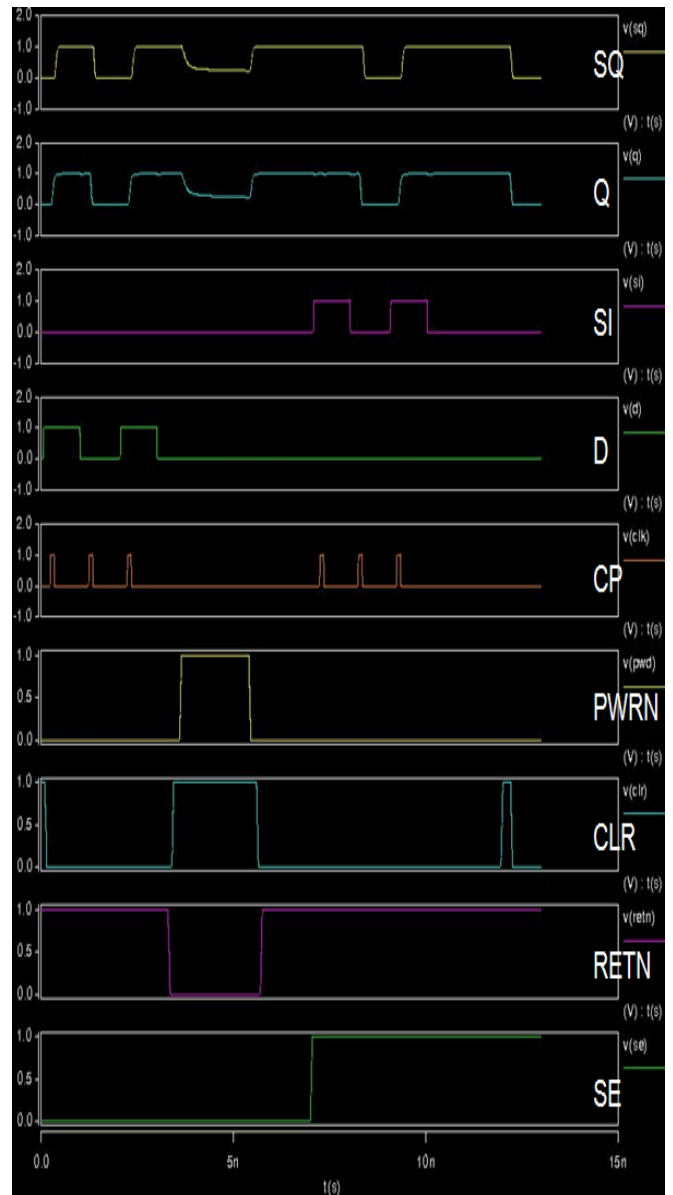


Fig. 4. Function, Scan, Reset, Retention/Restore modes simulation waveforms

V. AN AREA IMPROVEMENT DESIGN

After the proposed flop (LPTSPFFRR) was successfully verified, an improved design, named LPTSPFFRR1, has been developed based on LPTSPFFRR to further reduce area. The improvement was achieved by inserting the reset pull-down path between the clock and retention pass transistor P2 as shown in the circuit schematic in Fig. 5

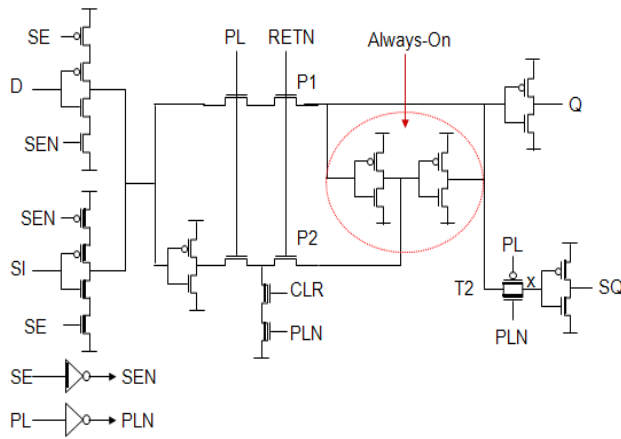


Fig. 5. Area reduction design of Low Power State Retention Testable Static Pulse-triggered Flip-flop

In this design, P2 functions as both the retention switch in the sleep mode and the path breaker of the asynchronous reset at wakeup. Consequently, the pull-down transistor P3 in LPTSPFFRR becomes not necessary and removed.

LPTSPFFRR1 was simulated by the same simulation deck of LPTSPFFRR. The results were very much similar to that of LPTSPFFRR simulation as expected.

VI. AREA AND POWER-DELAY EFFICIENCY COMPARISON

The proposed designs were evaluated and compared, in terms of power-delay efficiency and transistor size, with two leading-edge testable and resettable registers from two well-known low-power library vendors. One register is high performance retention register and another is optimized for low power.

The two proposed pulsed flops and the two selected production retention registers were simulated by a same simulation deck for fair comparisons. Two sets of simulations were performed.

In the first set of the simulation a single flop was simulated with an ideal pulsed clock at 1GHz frequency with 50ps pulse width and 50ps rise/fall slew time. This is to evaluate the flop without penalty from the pulse generator.

In the second simulation, an NAND gate pulse generator was used to drive a cluster of flops to evaluate the designs including the overhead of the pulse generator.

In both simulations, a 4x buffer was used to drive flop's inputs, and 10fF capacitor load (equivalent load of fanout 6 of 2x inverters) was added at the outputs of every flop.

The data were toggled at every clock cycle during power-delay measurement periods to simulate the worst-case operations. The simulations were run at typical corner, 1.0V VDD and 25°C temperature.

Table I shows normalized results of the single flop simulation with ideal clock sources.

Table I. Single flop power and delay comparison

Flip-flop	RMS power	Clk-Q delay	P*D product	Tran sizes
LPTSPFFRR1	1	1	1	1
LPTSPFFRR	1	1	1	1.03
High-speed flop	1.44	0.91	1.31	1.94
Low-power flop	0.98	1.34	1.31	1.56

The power and performance of the two proposed pulsed flops are almost identical due to the similar circuits. The area is about 3% improved in LPTSPFFRR1 than LPTSPFFRR.

Comparing with the high-performance register, the proposed pulsed flops are about the half size of the register. The power consumptions are about 44% less than the high-performance register. However, the clock-to-Q delay is about 9% longer. Nevertheless, the power-delay product of the proposed flops is 31% smaller than the high-performance register.

On the other hand, the low-power retention register is 2% less power consuming. However, it is 34% slower than the proposed pulse flops.

It is interesting to notice that the power-delay product of the high-performance and low-power retention registers is about the same which indicates that both registers have been optimized for optimal power-delay product, with one trading off power for speed and another is the other way around.

The area reduction of the proposed designs is significant. They are about the half size of the high-speed retention flop. The low-power retention register is a compacted design yet still 56% larger than the proposed designs.

The pulsed flops require a pulse generator to convert 50/50 duty cycle clock into a pulse sequence. This introduces power overhead in the design. To evaluate overall power-delay efficiency considering the power overhead of the pulse generator, a set of clusters of 12 to 60 pulsed flops were built. Each cluster is driven by a pulse generator through a clock buffer tree. Although sophisticated clock pulse generators could reduce power overhead, a simple NAND-style pulse generator was used in the evaluations due to constraints in project scope and effort. To make fair evaluations, the two leading-edge retention registers were also clustered in the same way and driven by the same size clock buffer tree of that used in the proposed flop simulation.

The HSPICE simulation results of the clusters in terms of the power-delay product are shown in Fig. 6.

The power overhead of the NAND-style pulse generator is considerable. As the result, the overall power-delay product improvement of the proposed pulsed retention flops in the clustered designs over the two leading-edge retention flops becomes less significant. From size of 48 flops, the cluster overall power-delay product improvement is about 9%. Nevertheless, the area efficiency of the proposed retention

pulsed flops remain significantly higher and leakage power is lower than the two leading-edge retention flops in the clustered designs.

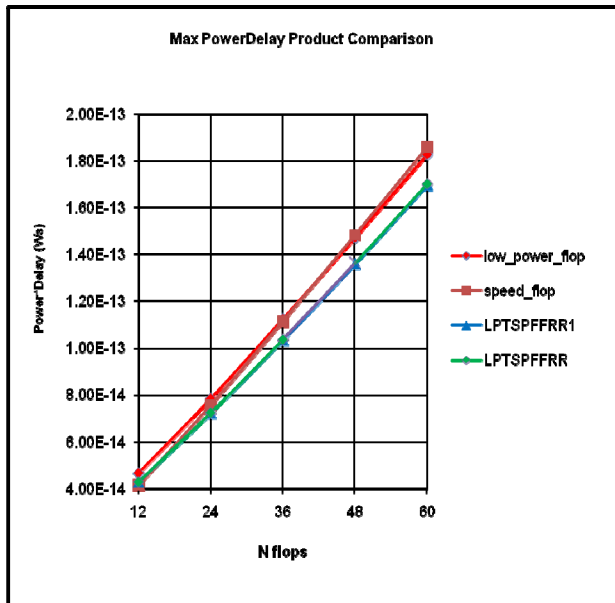


Fig. 6. Power-Delay product comparison of clusters of flops

VII. SUMMARY

Two area and power-delay efficient state retention pulsed flops with test and reset capabilities have been developed. The flops implement a single clock pin for function and test mode clocking. This simplifies clock distributions and reduces power on the pulse clock tree compared with three clocks implemented in the scan-latch based retention pulsed flops reported in [24]. An efficient retention circuit has been designed in the proposed flops with very low overhead using a single retention control signal instead of dual “save” and “restore” signals used in the balloon-style retention flops. The reset circuit of the proposed flops is also minimized consisting of only two transistors. As the result, the proposed flops are significantly more area and power-delay efficient than existing retention flops. The scan mode of the proposed design is compactable to existing scan-based test design. Therefore, it allows for leveraging conventional DFT flow. The non-overlapping SQ output of the design enables back-to-back connections of the pulsed flops in scan chains without hold-fixing buffers. This design feature can also be leveraged in function mode to drive short logic paths without having hold violations.

The higher area and power-delay efficiency and the compatibility to conventional retention flops design flows make the proposed retention pulsed flops useful to those high performance and low power ASIC designs which require state retention for rapid wakeup and low area and power overhead in the retention flops.

REFERENCES

- [1] Siva G. Narendra, Anantha Chandrakasan, “Leakage in Nanometer CMOS Technologies”, 2006
- [2] Kaushik Roy, Saibal Mukhopadhyay, and Hamid Mahmoodi-meimand, “Leakage current mechanism and leakage reduction techniques in deep-submicrometer CMOS circuits”, Proc. IEEE Vol. 91, no. 2, Feb. 2003
- [3] Dongwoo Lee, David Blaauw, and Dennis Sylvester, “Gate oxide leakage current analysis and reduction for VLSI circuits”, - IEEE Trans. VLSI, Vol. 12, No. 2, Feb. 2004
- [4] Kiat-Seng Yeo and Kaushik Roy, “Low-voltage, low-power VLSI subsystems”, McGraw-Hill, 2005
- [5] Enrico Macii, “Ultra low-power electronics and design”, Kluwer Academic Pub. 2004
- [6] Jan M. Rabaey and Massoud Pedram, “Low power design methodologies”, Kluwer Academic Pub. 2002
- [7] K. Kumagai, et al., “A Novel Powering-down Scheme for Low Vt CMOS Circuits,” 1998 Symposium on VLSI Circuits Digest of Technical Papers, pp. 44-45, 1998.
- [8] H. Makino, et al., “An Auto-Backgate-Controlled MT-CMOS Circuit,” 1998 Symposium on VLSI Circuits Digest of Technical Papers, pp. 42-43, 1998
- [9] David Flynn, Michael Keating, Robert Aitken, Alan Gibbons and Kaijian Shi , “Low Power Methodology Manual for System-on-Chip Design”, Springer, 2007
- [10] Hamid Mahmoodi-Meimand and Kaushik roy, “Data-Retention Flip-flops for Power-Down Applications”, Proc. ISCAS, 2004
- [11] Stephan Henzler, Thomas Nirschl, et al., “Dynamic State-Retention FlipFlop for Fine-Grained Sleep-Transistor Scheme”, Proc. ESSCIRC, 2005
- [12] Lawrence Clark, Mohammed Kabir and Jonathan Knudsen, “A Low Standby Power Flip-flop with Reduced Circuits and Control Complexity”, Proc. CICC, 2007
- [13] Yuan, J., and Svesson, C.: ‘New single-clock CMOS latches and flipflops with improved speed and power savings’, IEEE J. Solid-State Circuits, 1997, 32, (1), pp. 62–69
- [14] M.-W. Phyu, W.-L.Goh and K.-S. Yeo, ”Low-power/high-performance explicit-pulsed flip-flop using static latch and dynamic pulse generator”, IEE Proc.-Circuits Devices Syst., Vol. 153, No. 3, June 2006
- [15] Ghadiri, A., and Mahmoodi, H.: ‘Dual-edge triggered static pulsed flip-flops’. 18th Int. Conf. on VLSI Design, India, Jan. 2005
- [16] N. Nedovic, V.G. Oklobdzija, "Hybrid Latch Flip-Flop with Flip-Flop with Improved Power Efficiency," IEEE Symp. on Integrated Circuits and Systems Design, PP. 211-215, Sep. 2000.
- [17] Don Douglas Josephson, S. Poehlman, V Govan, C. Mumford, “Test Methodology for the McKinley Processor”, Proc. ITC International test conference, 2001
- [18] Samuel D. Naffziger, Glenn Colon-Bonet, Timothy Fischer, Reid Riedlinger, Thomas J. Sullivan, and Tom Grutkowski, “The Implementation of the Itanium 2 Microprocessor”, IEEE J. SOLID-STATE CIRCUITS, Vol. 37, No. 11, November 2002
- [19] Samuel D. Naffziger, Glenn Colon-Bonet, Timothy Fischer, Reid Riedlinger, Thomas J. Sullivan, and Tom Grutkowski, “The Implementation of the Itanium 2 Microprocessor”, IEEE J. SOLID-STATE CIRCUITS, VOL. 37, NO. 11, NOVEMBER 2002
- [20] Tshanz, J., Narendra, S., Chen, Z.P., Broker, S., Sachdev, M., and De, V.: ‘Comparative delay and energy of single edge-triggered and dual edge-triggered pulsed flip-flops for high-performance microprocessors’. Proc. ISPLED, Aug. 2001
- [21] A. S. Seyedi, S. H. Rasouli, A. Amirabadi, and A. Afzali-Kusha, “Low Power Low Leakage Clock Gated Static Pulsed Flip-Flop”, Proc. ISCAS2006, 2006
- [22] Y. Matsuya, et al., “A 1-V high-speed MTCMOS circuit scheme for power-down application circuits,” IEEE Journal of Solid-State Circuits, Vol. 32, pp. 861-869, Jun 1997.
- [23] H. Akamatsu, et al., “A Low Power Data Holding Circuit with an Intermittent Power Supply scheme for sub-1V MT-CMOS LSI,” Symposium on VLSI Circuits Digest of Technical Papers, 1996.
- [24] Victor Zyuban and Stephen Kosonocky, “Low Power Integrated Scan-Retention Mechanism”, Proc. ISLPEP, 2002