Energy Efficient Synchronous Sequential Circuits Design Using Clock Gating

M.Anusha, K. Nirmala, T. Lakshmi Prasanna

Assistant professor, Department of Electronics and communication engineering, Sree Vahini institute of science and technology, Tiruvuru, Krishna District, AP, India.

Associate professor, Department of Electronics and communication engineering, Sree Vahini institute of science and technology, Tiruvuru, Krishna District, AP, India.

Assistant professor, Department of Electronics and communication engineering, Sree Vahini institute of science and technology, Tiruvuru, Krishna District, AP, India.

ABSTRACT: Pulsed latches are gaining increased visibility in low-power ASIC designs. They provide an alternative sequential element with high performance and low area and power consumption, taking advantage of both latch and flip-flop features. While the circuit reliability and robustness against different process, voltage, and temperature variations are considered as critical issues with current technologies, no significant reliability study was proposed for pulsed latch circuits. In this paper, we present a study on the effect of different PVT variations on the behaviour of pulsed latches, considering the effect on both the pulsar and the latch. In addition, two novel design approaches are presented to enhance the reliability of pulsed latch circuits, while keeping their main advantages of high performance, low power, and small area. The two proposed designs have negligible power overhead when running at nominal supply voltage, and they have higher yield per unit power when compared with the traditional design at different voltages and temperatures.

KEY WORDS: Pulsed latches, flip-flops, pulsed flip-flops, variability, process variation, voltage scaling, low power.

I. INTRODUCTION

FLIP-FLOPS are considered the most popular sequential elements used in conventional ASIC designs. This is mainly because of the simplicity of their timing model, which makes the design and timing verification processes much easier. Master-Slave Flip-Flops (MSFFs) are considered the most common and traditional implementations of flip-flops, due to its stable operation and its simple timing characteristics. However, the fact that the MSFF micro-architecture is usually built using two consecutive latches, it takes an appreciable portion of the clock period, power consumption, and area. In addition to the mentioned Overheads associated with MSFF, some additional margins, which can reach up to 15% (depending on the sign off methodology), are usually added to the nominal timing margins to ensure correct operation under different process, voltage, and temperature (PVT) variations.
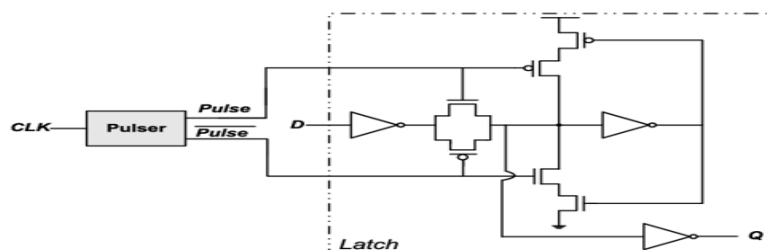

Fig. 1. Simple diagram of a traditional transmission gate pulsed latch.

To fill in the missing gap between MSFFs and latches, pulsed latches (sometimes called pulsed flip-flops) have been used in some high-performance designs. Pulsed latches (PLs) are latches driven by short pulses generated

from the normal clock signal using a pulse generator circuit called a *pulsar* as shown in Fig. 1. The pulsar can be either embedded in the latch, or can be separated as a standalone circuit as shown in Fig. 1.

If the latter approach is used, a single pulsar can be shared by more than one latch. Thus, it has the advantage of area and power consumption savings over the former approach and it is the focus of our discussion in this paper. In this paper, we are presenting variability analysis of one of the popular topologies of pulsed latches, Transmission Gate Pulsed Latch (TGPL), studying the effects of process, voltage, and temperature variations, as well as proposing design modifications that can help in decreasing the probability of circuit failure (i.e. enhancing pulsed latch reliability) at different supply voltage values. With the proposed approaches, pulsed latches present a formidable alternative to MSFFs, providing higher performance, lower area and power consumption, and higher reliability and robustness to different kinds of variations.

II. LITERATURE SURVEY

Pulsed latches have been always proposed to decrease power consumption and increase performance. In PLs with relatively wide pulse widths were used to allow cycle borrowing and tolerate any clock skew. In PLs were used as the main sequential elements to increase the performance of the Intel X Scale microprocessor without consuming high clock power.

Although the minimum pulse widths to ensure reliable operation across different PVT corners were used, delay buffers were still be needed to decrease the risk of hold time violations. However, some area and power overhead were presented due to buffer insertion. The paper showed that the proposed pulser is less affected by the clock rise time when compared with the traditional pulser at different supply voltages. However, the paper defined the failure criteria by the ability of the pulser to output a valid pulse, without quantifying the satisfaction of the pulse width for the needed latch transparency window in order to achieve successful writing. In addition, the study didn't quantify the effect of these variations on the design yield. Also, the studied voltage variations were limited to 10% only, while the temperature variations studied was only 20◦C around 85◦C. As shown from the previous studies, TGPL, which will be our focus in this paper, is one of the most attractive architecture for PL circuits. However, there are still some challenges in the TGPL design (PL in general) to ensure reliable operation under PVT variations. In addition, a more comprehensive study of the probability of failure based on both the pulser and the latch is still missing. Although the study and proposed architectures presented in this paper will focus on TGPL, the same approaches can be applied to any other PL topology.

III. EFFECT OF PVT VARIATIONS ON PULSED LATCHES

The operation of PLs is based on enabling the latch for a short time using a pulse generated by the pulse circuit. Hence, to study the effect of variation on PL operation, variation effects on both the latch write time and the pulse pulse width should be studied. The effect of process variations is carried out for each of the latch and the pulse independently and the same study is repeated for different voltage and temperature values of interest.

A. Process Variations

Due to the extreme miniaturization of device parameters in current and upcoming technology processes, even a small variation in the manufacturing process may cause parameter variations that can lead to a failed circuit operation. Thus, one of the significant challenges in the design phase is the ability to evaluate the effect of different sources of variations on the functionality of complex circuits and to provide circuit solutions to guarantee correct functionality under different sources of variations. Process dependent sources of variability such as effective length variation, oxide thickness variation, Line Edge Roughness (LER), and Random Dopant Fluctuation (RDF) (for planar MOSFETs) result in variations in the value of the threshold voltages of transistors, which in turn impact the timing and power of digital circuits. The threshold voltage variations due to RDF (which is usually the principal source of threshold voltage variations in planar MOSFETs) are considered as zero-mean Gaussian independent random variables with standard deviation denoted as σVth which is given by:

$$\sigma_{Vth} = \sigma_{Vtho}\sqrt{\frac{L_{min}W_{min}}{LW}} \qquad (1)$$

Where $\sigma Vtho$ is the σVth for minimum sized transistor and it is given by:

$$\sigma_{Vtho} = \frac{q T_{ox}}{\epsilon_{ox}} \sqrt{\frac{N_a W_d}{3 L_{min} W_{min}}} \qquad (2)$$

Where Na is the effective channel doping, Wd is the depletion region width, Tox is the oxide thickness, $Lmin$ and $Wmin$ are the minimum channel length and width, respectively. While the scaling down of CMOS technology reduces the nominal supply voltage, the threshold voltages are not scaled by the same factor, leading to a significant reduction of the transistor's available voltage headroom (the difference between the supply voltage and the threshold voltage). Hence, even any small variation in the transistor threshold voltage can lead to a significant degradation of the circuit behavior or can even cause complete circuit failure.
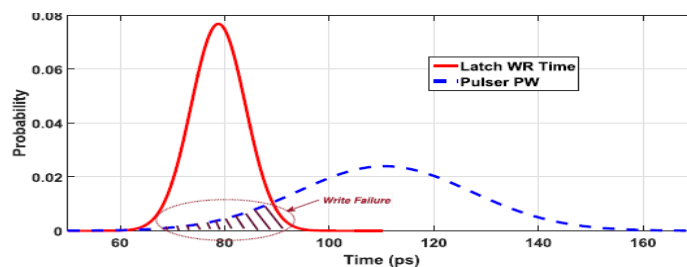

Fig. 2. Sample PDFs of the latch write time and the pulser pulse width showing the region of write failure.

Studying the effect of process variations on PLs includes studying the effect of variations on both the latch write time and the pulser pulse width. As shown in Fig. 2, this is represented by the probability distribution functions (PDFs) of both the write time (Latch WR Time) calculated as the CLK-to-Q delay and the pulse width (Pulser PW).

To ensure correct write operation, the pulse width should be larger than the required transparent window for the latch (i.e. time needed to capture the input data and pass it through the internal nodes to the storing cross coupled inverters). The area under the intersection between the two PDFs represents the failure of write operation, since this is the region where there is a high probability that the pulse width will be smaller than the time needed by the latch. Alternatively, knowing the information about the distribution of the latch write time and for easiness of timing analysis, a maximum value for latch write time can be calculated for certain sigma value of the designer choice.

In this case, the probability of write failure can be calculated as the probability of having the width of the pulser output smaller than this desired maximum value. In both cases, depending on the target yield, the designer can determine the minimum acceptable value for circuit failure, and hence, the transistors' dimension can be adjusted to reach the target yield.
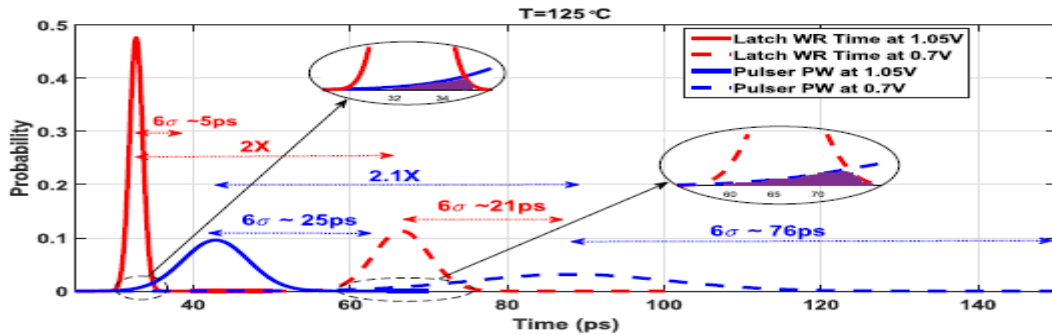
Fig. 3.   The effect of voltage scaling on the distributions of the latch WR time and the pulser PW at 125°C.

B. Voltage Scaling

Voltage scaling is a popular run-time technique used for reducing the power consumption of circuits. It significantly decreases both dynamic power (with its two components of switching power and internal power) and leakage power.

On the other hand, the ability to reduce the operating supply voltage is limited by a minimum value determined usually by some timing constrains (critical path delay as an example), in addition to some margins for the PVT variations, and usually adding a margin for aging effects. As the supply voltage is scaled down, the available voltage headroom decreases further and the transistors become more sensitive to any variations.

The effect of voltage scaling is naturally associated with the increase of timing delays for different circuit components. While this can be handled at design time for several circuit components, the case may not be as easy for PLs. Since PL operation depends on two different components (pulser and latches) of different micro architectures, the timing of each of them is affected differently. As shown in Fig. 3, voltage scaling affects the probability distribution of the pulser and the latch differently.

As shown in Fig. 4, the probability of write failure for a PL can increase by up to two order of magnitude when the supply voltage is scaled down by around 30%. Even if the PL circuit is designed to operate reliably at an intermediate supply voltage (0.9V as an example), the reliability will still significantly degrade at lower voltages, especially at low operating temperatures. One possible solution is to design the PL circuit to operate with the needed level of reliability at the lowest possible operating voltage. Since chips usually operate at different supply voltages with different operating modes, when pulsed latches are operating at a voltage higher than that minimum value, they will be operating with extra timing margin (the pulse width will be larger than the needed width to achieve the required level of reliability).
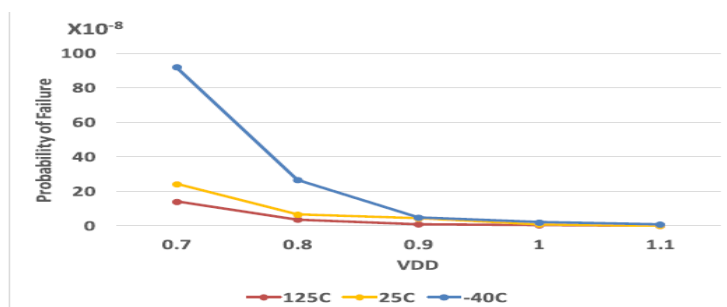

Fig. 4.   The probability of failure of a traditional pulsed latch designed at nominal supply voltage at different supply voltages and temperatures.

C. Temperature Effect

Studying the effect of temperature variation on the design is very important. Not only does the variation in temperature affect leakage power and performance, but it also affects the probability of having an error during circuit operation, as well as impacting the life span of different chip parts. Factors such as the increase of leakage power with technology process scaling, the non-equivalent down scaling of the supply voltage when compared to geometry scaling, and the increase in the dynamic power associated with the increase in performance

The study done in this paper shows that both circuits become more sensitive to process variation with the decrease in temperature.

IV. SEQUENTIAL CIRCUITS DESIGN USING CLOCK GATING USING MENTOR GRAPHICS

The Mentor Graphics software package consists of a large number of executable files, documents, libraries, and other components. The locations of these files vary from system to system, so it is necessary to incorporate some mechanism for handling the differences that naturally occur between the installations at different sites. Many of these details are handled using start-up scripts and environment variables. Although these start-up scripts simplify the use of the software for the end user, it may be necessary for the user to do some minor editing of start-up files before the software can be used.

As described in the previous section, it is not easy to design a non-configurable pulsed latch circuit that can operate with just the needed timing margins at different supply voltages in the presence of process and temperature variations, while keeping the needed level of reliability. To be able to reach the needed reliability level, the pulser circuit should be reconfigured at run time to generate an output pulse whose width can be controlled based on the operating condition. In this section, two design approaches are proposed. Both approaches depend on controlling the delay path (the delay unit and its following inverter) of the pulser circuit by using an external control signal (CTRL) to generate a controllable pulse width. The first approach considers splitting the supply rail of the pulser circuit, and applying an additional controllable level of voltage scaling on the delay path when needed. The second approach relies on using multiple delay units in the pulser circuit and choosing a certain delay unit at run-time according to the operating condition. Detailed discussions of the two approaches are presented in the next two subsections.
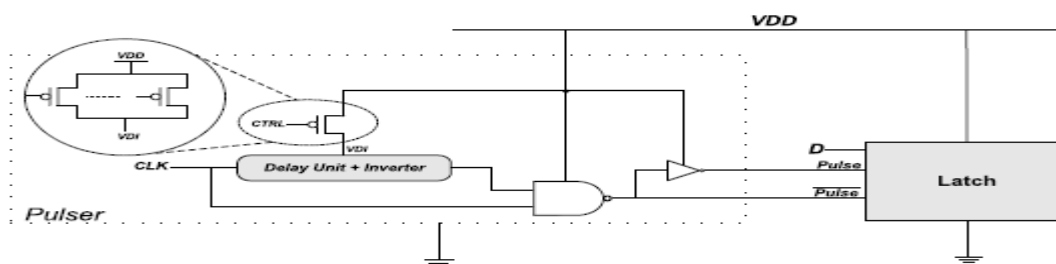

Fig. 5. The switch-based pulser design.

A. First Approach

This approach is based on using a virtual supply rail for the delay path of the pulser, driven from the main supply rail used for the rest of the pulser circuit and the latches. This can be accomplished using header PMOS switches for the delay path of the pulser circuit, similar to the local power gating topology as shown in Fig. 5, where turning off some of these switches will result in lowering the supply voltage of the delay path.

Since this delay path is the main part of the circuit that control the width of the generated pulse, controlling the supply voltage of this path will result in controlling the output pulse width. Separate control signals can be used for different switches, where at least one of these switches must be always turned on (i.e, the gate of this PMOS switch should be tied to the ground) giving the maximum output pulse width, while the other switches can be turned on or off to achieve the required narrowing of the pulse width. The number of these parallel switches and their sizes will depend on the number and values of the virtual supply voltage levels, which corresponds to the needed pulse widths to achieve the target reliability level at different operating conditions.

When scaling down the supply voltage, the needed margin for variations in the latch writes time increases. In addition, the remaining pulser circuit (the NAND gate and the output inverter) will act as a voltage level shifter, driving the latches by the same voltage level as their supply voltage.
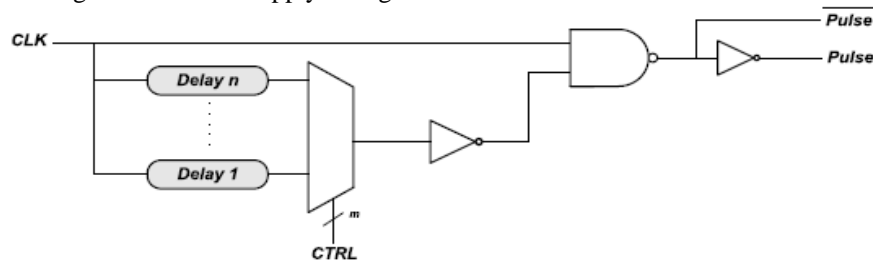

Fig. 6. The MUX-based pulser design.

B. Second Approach

Implementing multiple delay units with different delays can help in generating pulses with different widths. One important design consideration is the ability to choose between these different units post silicon or at run-time.

The second proposed pulser design is shown in Fig. 6. Each delay unit represents a buffer chain that can be implemented in different ways. It can be as simple as a very small delay unit (i.e., just a wire) and up to multiple even number of inverters of different inverter sizes and/or numbers.

The output of the multiplexer is used to drive an odd number of inverters, whose final output is connected to the NAND gate. By selecting a longer delay chain, the latch transparency window can be increased at run time, which is required when scaling down the supply voltage. The shortest delay unit is designed such that, when operating at a nominal supply voltage, the circuit is verified to run with very low probability of failure in the presence of different process and temperature variations. The rest of the delay units are designed depending on the number and values of the supply voltage scaling levels.

V. EXPERIMENTAL RESULTS

To verify the proposed approaches, test circuits of 16-bit register were examined and three implementation choices were compared. The three implementations consist of a single pulser driving sixteen identical latches similar to that shown in Fig. 1. The first choice is the implementation using the traditional non-configurable pulser shown in Fig. 6. The pulser was designed at nominal supply voltage to ensure the required reliability level. The second and the third choices are the two proposed pulser implementations, also driving sixteen identical latches. The effect of voltage scaling of one scaling level was applied on all circuits. An extreme value of voltage scaling which is usually around 30% reduction from nominal supply value was used to show the effectiveness of the proposed approaches. The same approaches can be easily extended to any other scaling values.

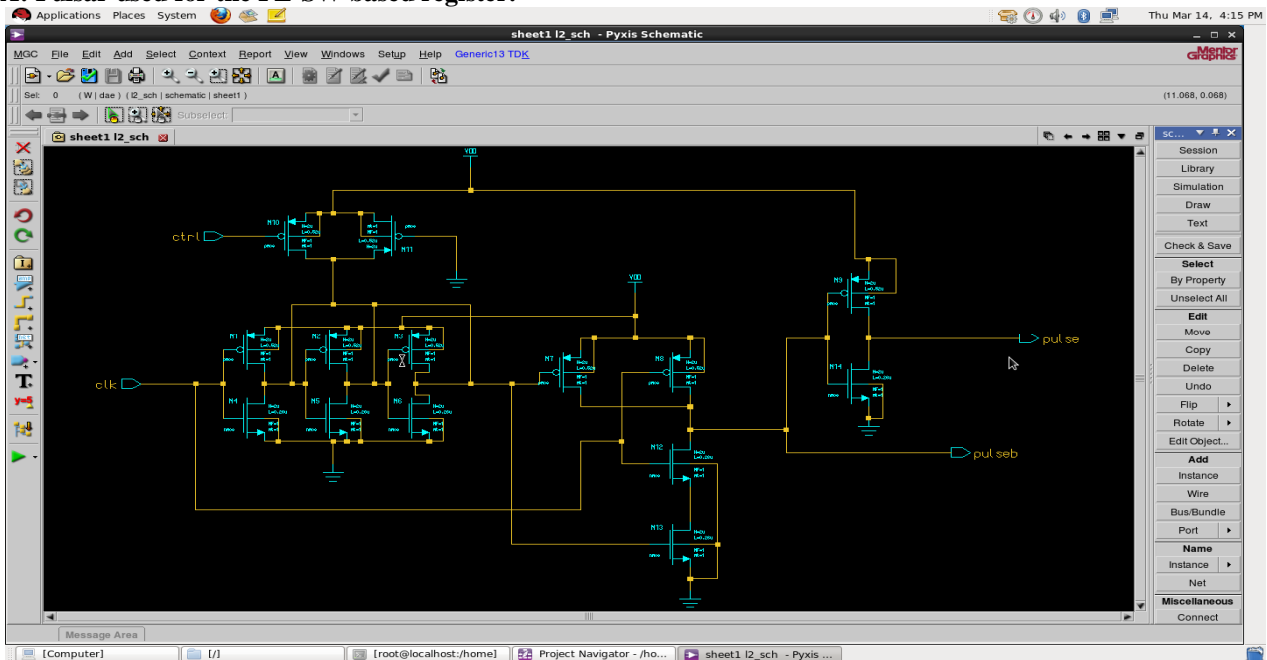**A: Pulsar used for the PL-SW based register:-**

Fig: Pulsar used for the PL-SW based register

A.1: Simulated results of PL-SW based register:-

Fig: Simulated results of PL-SW based register
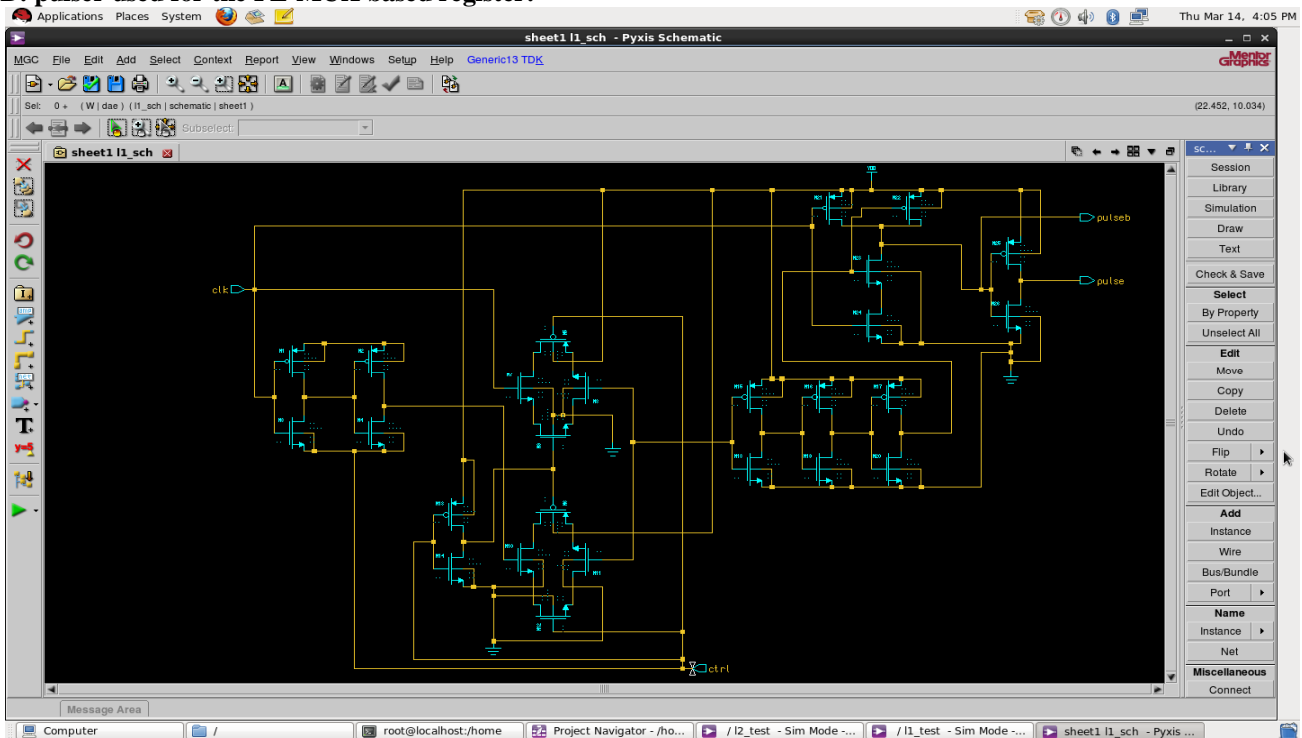
B: pulser used for the PL-MUX based register:-

Fig: pulser used for the PL-MUX based register

B.1: Simulated results of pulser used for the PL-MUX based register:-

Fig: Simulated results of pulser used for the PL-MUX based register

Hence, for designs with few voltage scaling levels, the PL-MUX design can be preferable over the PL-SW one, as it is easier in design, generates more precise pulse widths, and its overheads (power and area) are reasonable. On the other hand, for designs with large number of voltage scaling levels, the PLSW design is preferred, as the area and power overheads of the PL-MUX design will be significant.

VI. CONCLUSION

In this paper, an analysis of the effect of PVT variations on the pulsed latch performance was presented. The analysis considered both the pulser and the latch to evaluate the reliability of the entire pulsed latch circuit. In addition, the benefits of having a reconfigurable pulsed latch circuit were discussed. Two novel modifications to add the reconfiguration ability to TGPL circuits were proposed. The benefits of using the proposed design approaches in enhancing the robustness of pulsed latch circuits at different supply voltages were demonstrated using 16-bit registers. Both proposed approaches were able to ensure reliable operation of the pulsed latch-based register under different supply voltages in the presence of process and temperature variations, without any unnecessary timing overhead. Both approaches have a very small area overhead of around 3% or less. In addition, the power overhead of both approaches is minimal when compared to the traditional pulsed latch based register at the same reliability level. Both approaches are easily scalable to cover different levels of voltage scaling. In addition, they can be applied to any other pulsed latches topology that depends on a delay path to generate the output pulse.

REFERENCES

[1] D.Chinnery and K.Keutzer, *Closing the Gap Between ASIC & Custom: Tools and Techniques for High-Performance ASIC Design*. Norwell, MA, USA: Kluwer, 2002.

[2] S.Paik, G.-J.Nam, and Y.Shin, "Implementation of pulsed-latch, and, pulsed-register circuits to minimize clocking power," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2011, pp. 640–646.

[3] Y.Shin and S.Paik, "Pulsed-latch circuits: A new dimension in ASIC design," *IEEE Des. Test Comput.*, vol.28, no.6, pp.50–57, Nov./Dec.2011.

[4] M.A.Alam, K.Roy, and C.Augustine, "Reliability-and-process-variation aware design of integrated circuits—A broader perspective," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Apr.2011, pp.4A.1.1–4A.1.11.

[5] E.Consoli, G.Palumbo, J.M.Rabaey, and M.Alioto, "Novel class of energy-efficient very high-speed conditional push–pull pulsed latches," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol.22, no.7, pp.1593–1605, Jul.2014.

[6] J.Warnock *et al.*, "Circuit and physical design of the zEnterprise EC12 microprocessor chips and multi-chip module," *IEEE J. Solid-State Circuits*, vol.49, no.1, pp.9–18, Jan.2014.

[7] T.Baumann *et al.*, "Performance improvement of embedded low-power microprocessor cores by selective flip flop replacement," in *Proc. 33rd Eur. Solid State Circuits Conf. (ESSCIRC)*, Sep.2007, pp.308–311.

[8] L.T.Clark *et al.*, "An embedded 32-b microprocessor core for low-power and high-performance applications," *IEEE J. Solid-State Circuits*, vol.36, no.11, pp.1599–1608, Nov.2001.

[9] M.Alioto, E.Consoli, and G.Palumbo, "Analysis and comparison in the energy-delay-area domain of nanometer CMOS flip-flops: Part I—Methodology and design strategies," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol.19, no.5, pp.725–736, May2011