SGERC: a self-gated timing error resilient cluster of sequential cells for wide-voltage processor

Taotao Zhu[1], Xiaoyan Xiang[2a], Chen Chen[2], and Jianyi Meng[2]

[1] *Institute of VLSI Design, Zhejiang University, Hangzhou, 310027 China*

[2] *State Key Laboratory of ASIC and System, Fudan University, Shanghai, 201203 China*

a) *xiangxy@fudan.edu.cn*

Abstract: This paper presents a self-gated error resilient cluster of sequential cells (SGERC) to sample the critical data in wide-voltage operation for EDAC system. SGERC introduces latch-based clock gating technique to error resilient circuits and proposes a customized clock gate which has the ability of timing error self-correction with only two additional transistors added for the first time. Further, it totally eliminates the timing error detection circuits required by each critical register before and utilizes the data-driven clock gating circuits to generate timing error information. Simulation results show that SGERC design achieves 58.3% energy efficiency improvement compared with the baseline design and 19.4% over the latest EDAC design.

Keywords: error resilient, clock gating, wide voltage, low power

Classification: Integrated circuits

References

[1] S. Das, *et al.*: "RazorII: in situ error detection and correction for PVT and SER tolerance," IEEE J. Solid-State Circuits **44** (2009) 32 (DOI: 10.1109/JSSC.2008.2007145).

[2] S. Wimer and I. Koren: "The optimal fan-out of clock network for power minimization by adaptive gating," IEEE Trans. VLSI Syst. **20** (2012) 1772 (DOI: 10.1109/TVLSI.2011.2162861).

[3] K. A. Bowman, *et al.*: "Energy-efficient and metastability-immune resilient circuits for dynamic variation tolerance," IEEE J. Solid-State Circuits **44** (2009) 49 (DOI: 10.1109/JSSC.2008.2007148).

[4] S. Valadimas, *et al.*: "Effective timing error tolerance in flip-flop based core designs," J. Electron. Test. **29** (2013) 795 (DOI: 10.1007/s10836-013-5419-3).

[5] Y. Zhang, *et al.*: "8.8 iRazor: 3-transistor current-based error detection and correction in an ARM Cortex-R4 processor," IEEE International Solid-State Circuits Conference (ISSCC) (2016) (DOI: 10.1109/ISSCC.2016.7417956).

[6] I. Kwon, *et al.*: "Razor-Lite: A light-weight register for error detection by observing virtual supply rails," IEEE J. Solid-State Circuits **49** (2014) 2054 (DOI: 10.1109/JSSC.2014.2328658).

[7] P. Gupta, *et al.*: "Underdesigned and opportunistic computing in presence of

hardware variability," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **32** (2013) 8 (DOI: 10.1109/TCAD.2012.2223467).

[8] S. Roy, *et al.*: "Clock tree resynthesis for multi-corner multi-mode timing closure," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **34** (2015) 589 (DOI: 10.1109/TCAD.2015.2394310).

[9] S. Wimer, *et al.*: "The optimal fan-out of clock network for power minimization by adaptive gating," IEEE Trans. VLSI Syst. **20** (2012) 1772 (DOI: 10.1109/TVLSI.2011.2162861).

[10] W. Shan, *et al.*: "Timing monitoring paths selection for wide voltage IC," IEICE Electron. Express **13** (2016) 20160095 (DOI: 10.1587/elex.13. 20160095).

[11] D. Bull, *et al.*: "A power-efficient 32 bit ARM processor using timing-error detection and correction for transient-error tolerance and adaptation to PVT variation," IEEE J. Solid-State Circuits **46** (2011) 18 (DOI: 10.1109/JSSC. 2010.2079410).

[12] S. Kim and M. Seok: "Variation-tolerant, ultra-low-voltage microprocessor with a low-overhead, within-a-cycle in-situ timing-error detection and correction technique," IEEE J. Solid-State Circuits **50** (2015) 1478 (DOI: 10. 1109/JSSC.2015.2418713).

1   Introduction

With the increasing demand for power saving, near-threshold voltage computing occupies an important position. However, it is vulnerable to process, supply voltage, temperature, and aging (PVTA) variations, which is addressed by operating the processor at conservative voltage and frequency points [1]. Further, large safety margins incur great losses in area, energy and performance. To settle these issues, timing error resilient technique is proposed to protect the circuits from variations and eliminate the excessive margins. It uses timing error detection and correction (EDAC) mechanism to monitor the timing error violation and correct it when error occurs.

But for EDAC system, its clock network also faces the timing error problem if it adopts clock gates to decrease its dynamic power. Typically the clock tree is responsible for up to 50% of the total dynamic power consumption [2] and clock gating is a predominant technique to help preventing unnecessary switching of clock signals. Meanwhile, clock gate usually adopts the latch-based style to avert glitches on its enable signal from propagating to registers' input clock. So it must meet the timing constraints such as setup time check. And if it faces a timing violation, a cluster of registers it serves may keep the false data. For now, few papers target at this and explore the working characteristics of clock gating for error resilient system.

Furthermore, researchers now still focus at the single EDAC cell and make great efforts to decrease its area and power cost, which has arrived at the bottleneck. [3] proposes double-sampling design to detect timing error by adding additional memory elements and [4] is based on the transition detector. Recently, iRazor [5] uses only three-transistor current-sensing circuit to detect timing violation and adopts error mask technique to recover with one-cycle penalty. For timing error

correction, error mask technique has a better performance and can be realized without the modification of processor architecture compared with the previous replay mechanism [6]. So it's difficult to optimize EDAC design further by enhancing single cell.

In this paper, we introduce the clock gating technique to EDAC system and totally eliminate the previous error detection circuits required by each critical register. A self-gated timing error resilient cluster of sequential cells (SGERC) is proposed, which only relies on the data-driven clock gating network to flag the timing violation information. And with only two transistors added, it is the first time that EDAC system employs an error resilient clock gate called TESCG to restore the clock signals when the clock network faces a timing violation. Furthermore, an automated insertion methodology is presented to realize it into the commercial design.

2 Preliminaries

If error resilient systems employ clock gating technique to decrease clock tree power, it must add EDAC mechanism for clock gates. Although removing all the clock gates in the critical path and keeping the non-critical ones is an alternative way to solve this problem, it will cut down the benefits from clock gating greatly. Because the clock gate which occurs in the critical path takes a large proportion. For example, the critical clock gates account for 68% of the total ones based on the post P&R design of the commercial C-SKY CK802 micro-processor. And if its endpoint slack of timing path is less than 20% of the clock period, this cell is referred as critical clock gate or register.

The reason for the above phenomenon can be concluded into two aspects. First, the enable signal of clock gate derives from the input data of registers and the critical register has a large presence. [6, 7] show the critical wall phenomenon of register in different architecture and the critical register in razor lite takes 55%. Second, the commercial P&R tool tends to place clock gates near the clock tree root to control more cells during the clock tree synthesis [8], which makes their timing constraints much more tighter.

If a design intends to introduce error resilient clock gate, one way is to add the same EDAC circuits for clock gate as the register's such as transition detector. However, this approach takes a great hardware cost and energy consumption. And it doesn't utilize the relationship between clock gating and EDAC technique. The clock gating circuit must monitor whether the input data of sampling cells change before the rising edge of clock signal while EDAC detects the late arriving of data when clock is high including its setup time. Since they both intend to watch the data changing in different period of one clock cycle, circuits sharing can be chosen to solve this problem.

3 Proposed circuit and insertion methodology

3.1 Circuit structure and working mechanism

SGERC is inserted into the critical path to replace the original critical registers and protects circuit from PVTA variations. As shown in Fig. 1, the cluster involves k

positive latches and one TESCG. In error-free operation, SGERC behaves as normal sampling cells to store the pipeline data and gates the clock by itself as needed. When timing error occurs, it generates timing violation information and performs the self-correction by all the cells in the cluster including TESCG current cycle.
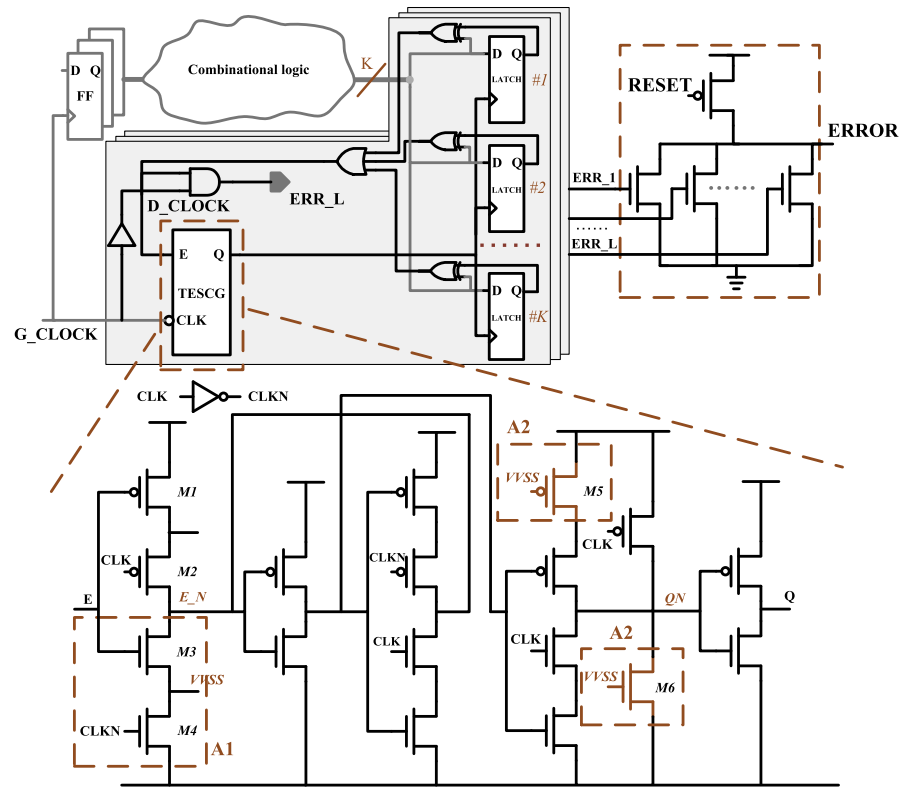

Fig. 1. SGERC's schematic including k sampling cells and a TESCG with data-driven clock gating.

For timing error detection, the error flag ERR_L is mainly generated from the data-driven clock gating circuits. Since the commonly used synthesis-based gating still leaves a large amount of redundant clock pulses, the data-driven method [9] is being researched to overcome this problem. SGERC groups critical latches whose switching activities are highly correlated and derives a joint enabling signal for clock gate by k XOR gates ORed together. Moreover, if this enabling signal rises during the high level of clock signal, it means that TESCG faces a timing violation. Meanwhile, through circuits sharing, this mechanism can also be used as the error indication of sampling cells in the cluster. Because the XOR gate will generate high output once any input data of these cells arrive late. So through the AND gate with the delayed clock signal, the cluster can indicate the timing error for all cells included.

As for error correction, TESCG adopts a negative latch to gate the clock and can perform self-correction when clock is high as the same as negative latches in the cluster. It is modified from a commercial latch-based integrated clock gate of SMIC 40 nm LL technology library. And it adds two additional transistors in area A2 and changes the transistor order in area A1 as pointed out in the orange

dotted box. Through the current change of VVSS when the data E arrives late as razor-lite [6], TESCG can turn on the newly added transistors to provide the appropriate clock signal.
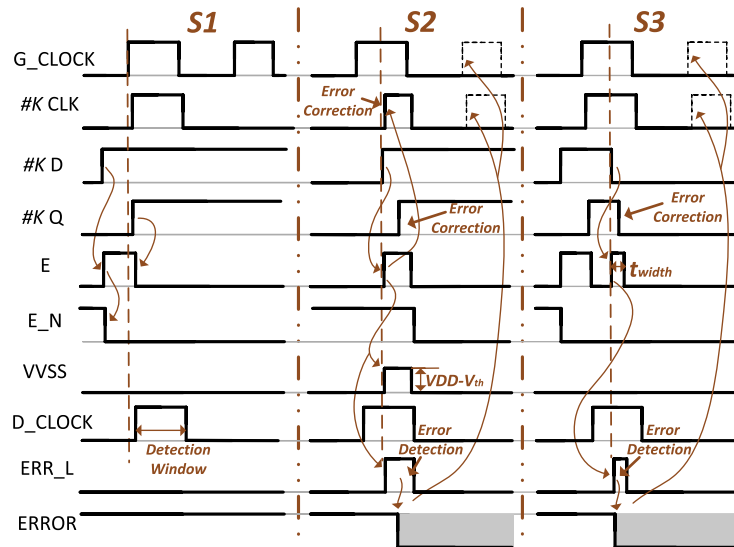

Fig. 2. The waveform of working mechanism including both the error-free and EDAC operation.

The detailed working mechanism is discussed with the conceptual timing diagrams in Fig. 2. When the system works with no timing error (*S1*), the input data of *#k* latch changes and it needs to be updated with new value. Its corresponding XOR gate generates a rise signal and TESCG receives the enable request and provides a useful clock signal for *#k* latch to sample data. In the next cycle, the input data remains unchanged, its clock signal is gated. For situation *S2*, its input data arrives late and the enable signal of TESCG changes to high when its clock signal G_CLOCK is high. The timing error signal ERR_L rises and TESCG performs self-correction operation. The inner node VVSS of TESCG charges to be high enough so that transistor M6 turns on. In order to accelerate the discharge of QN, transistor M5 turns off to interrupt the influence from the cross-coupled inverters. And the output of TESCG turns to high and provides a useful clock signal for latch sampling data. As a result, all the sequential cells in the cluster accomplish self-correction operation. Situation *S3* is a corner case when the input data changes twice or more during the one-cycle calculation. First, the intermediate value of D enables the clock gate and the latch samples wrong data. Second, the right input data arrives and XOR gate generates a pulse again which makes the error signal ERR_L enable. Since the latch can correct the data by timing borrow, the cluster can maintain the right operation with clock signal not gated. Besides, the high level width of E t_{width} should ensure its recognition by the following XOR gate, which is verified by Monte Carlo simulation at different corners. If it fails in some kinds of technology library, the design can also add delay buffer to increase t_{width} and it has no bad effect to clock gating network or EDAC ability.

Since timing error mask technique needs to borrow timing from the next stage in the pipeline, the next stage is very likely to face the timing violation if they are

the cascaded paths. To eliminate this bad effect, we choose to gate the whole pipeline one cycle to give the processor more time to calculate. For one-cycle error correction technique, it requires all the error information to be gathered current cycle and the error collection path may face a tight timing constraints. In SGERC design, it generates fewer error output signals. And the more cells clustered, the less timing risk undertaken. Moreover, to accelerate this process further, we use the dynamic OR latch logic in Fig. 1. Once the error of any cluster occurs, it discharges to zero and keeps the error information until the reset signal RESET enables.
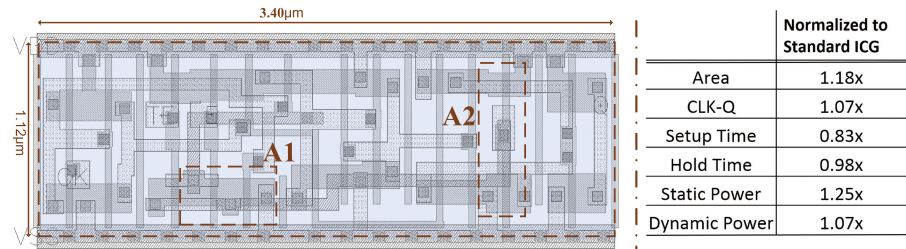
	Normalized to Standard ICG
Area	1.18x
CLK-Q	1.07x
Setup Time	0.83x
Hold Time	0.98x
Static Power	1.25x
Dynamic Power	1.07x

Fig. 3. Layout and circuit characteristics of TESCG.

The layout of TESCG and its circuit characteristics normalized to the unmodified ICG are shown in Fig. 3. Since the max value of VVSS is VDD-V_{th}, it needs to use skewed transistors to accelerate the signal transferring. After the modification at area A1 and A2, the cell enlarges by 18%. Then we extract the layout-parasite parameters and use SPICE to perform post-layout simulation. Compared to the original one, its CLK-Q delay increases by 7%. Moreover, because TESCG can correct the timing error when E rises to high, the fall setup time is given. It lowers to 83% and the hold time is 98% since node VVSS provides extra capacitance to improve the drive strength of the first tri-state inverter. With two additional transistors added, the static power enlarges by 25% and dynamic power 7%.

3.2 Timing constraints analysis

The timing constraints analysis of SGERC includes both the critical sampling latches and TESCG. The detection window T_{DW} in Fig. 4 is defined as a time interval when the system must monitor the timing violation and boot the error correction mechanism as needed.
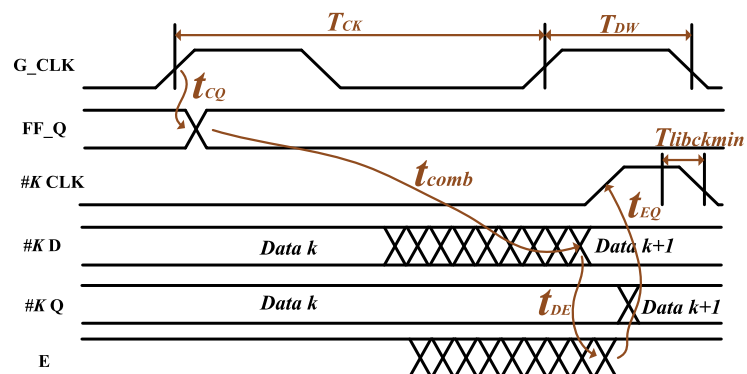
Fig. 4. Timing path analysis of SGERC design.

Using the notations t_{CQ}, t_{DE}, t_{EQ}, t_{ctd} for the propagation delay of CLK-Q of sampling latch, input data D to the enable pin E and E-Q of TESCG and the clock tree, respectively, t_{comb} for the combinational logic delay between two pipeline stages and T_{CK} for the system clock period, the following constraints must be satisfied for TESCG:

$$t_{CQ} + t_{comb} + t_{DE} + t_{setup\_CG} < T_{CK} + t_{vw} \qquad (1)$$

$$t_{vw} = T_{DW} - T_{libckmin} - t_{EQ} + t_{CQf} \qquad (2)$$

where $T_{libckmin}$ is the minimum clock width requirement for the sampling latch in the technology library and t_{CQf} is the CLK-Q delay of TESCG when clock falls. Equation (2) shows that the valid timing window size t_{vw} of TESCG is less than T_{DW} because it must provide a useful clock for sampling latches. Moreover, due to the short path problem, the hold time check of TESCG is defined as:

$$t_{CQ} + t_{comb} + t_{DE} > t_{vw} \qquad (3)$$

The timing constraints for the sampling latches are nearly the same with other EDAC technique. Due to the timing borrow ability, the setup time can have an additional timing period T_{DW}.

To evaluate the EDAC ability of SGERC in near-threshold voltage operation, its response time of error self-correction is evaluated. Taking the variations into consideration, a 4 K-point Monte Carlo simulation is performed to show the propagation delay from the input data D to the output right data Q for situation *S2* at 0.6 V. As shown in Fig. 5, its max value is almost 13 ns while the mean value is 1 ns, which is enough for a system working at 20 MHZ (@0.6 V) such as IOT processor when T_{DW} takes 25%. To accelerate this procedure further, we have replaced critical cells with LVT circuits.
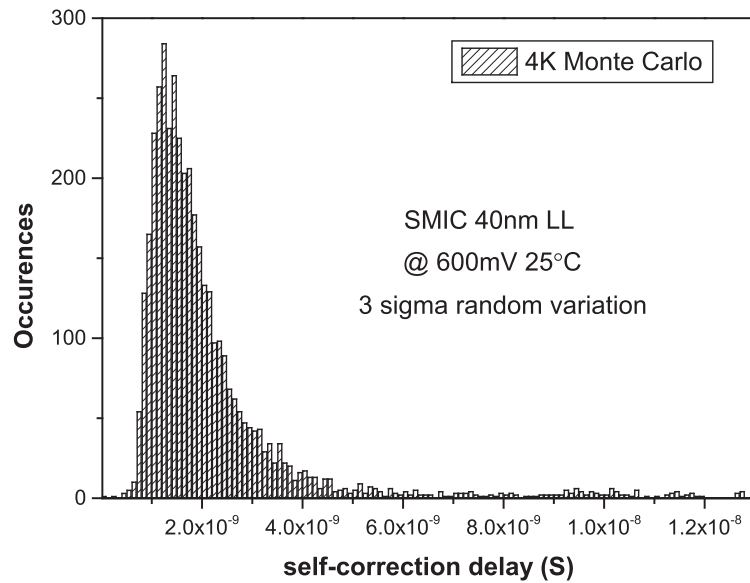

Fig. 5. The self-correction delay of SGERC by 4k-point Monte Carlo simulation.

3.3 Insertion methodology

An automated insertion algorithm and flow are proposed to embed SGERC into commercial design. Since it develops the gating and error detection logic shared by k latches, it may increase the amount of redundant clock pulses for each cell in the cluster.

Let the average toggling probability of a sampling cell be denoted by p $(0 < p < 1)$. Based on p and cell capacitance in the cluster, group size k for the maximized power savings is derived from [9] and it's the solution of

$$(1 - p)^k ln(1 - p)(C_{SL} + C_W) + (C_{TESCG} + C_{AND})/k^2 = 0 \qquad (4)$$

where C_{SL} is the clock input capacitance of sampling latch, C_W is the unit-size wire capacitance, C_{TESCG} is the TESCG capacitance and C_{AND} is the input capacitance of AND gate in the clock tree. We use SPICE to extract the parameters in Equation (4) and finally get the optimal cluster size k on different toggling probability as shown in Table I.

Table I. Optimal cluster size on toggling probability

p	0.01	0.02	0.04	0.06	0.1
k	14	10	8	7	6

After calculating the optimal group size k, it needs to group appropriate critical sampling cells into k-size sets. We propose the Cells Clustering Algorithm to operate the grouping operation as shown in Table II. Initially, the critical cells are sorted as ascending order of toggling probability. Then it intends to cluster k cells in a group using the optimal k value of the first cell in the set $\{C_i\}$. Further, we introduce the cell position in the layout to make the grouping more accurately. For cells with same k, it intends to choose those whose sum of distance is the minimum value.

The automated design flow is discussed below. First, we get the critical sequential cells which are reported by the static timing analysis. Second, toggling probabilities of these cells are estimated through running a package of benchmark. Third, the preliminary preferred locations of FFs in the layout are evaluated by the placement tool. Fourth, an automated tool insert the SGERC to the design according to Cells Clustering Algorithm. Finally, the whole physical design flow is operated and timing convergence in different voltage operation is achieved as [10].

4 Experimental evaluation

4.1 SGERC processor implementation

The proposed SGERC approach is implemented in a commercial C-SKY CK802 processor which has a 3-stage pipeline with a certain level of performance (1DMIPS/MHZ). The physical design is based on SMIC 40 nm LL process and the circuit is expected to work at 0.6~1.1 V with DVFS system.

First, a baseline design without EDAC circuits is introduced in the traditional design flow with margins added. We add clock gating technique for its lowest

Table II. Cells Clustering Algorithm

Input: toggling probability set $\{p_i\}$, locations of cells in the layout
Output: cluster sets
Algorithm:
1. Sort n cells $\{C_i\}$ such that $p_1 \leq p_2 \leq \cdots \leq p_n$;
2. **while** $\{C_i\}$ **do**
3.   Decide the optimal k of C_1, based on Equation (4);
4.   **if** $k < 2$ **then**
5.   **break**;
6.   **end**
7.   **foreach** $\{C_i\}$ **do**
8.   Count cell number j with the same k;
9.   **end**
10.   **if** $j \geq k$ **then**
11.   Group k cells in a cluster based on minimum distance;
12.   **else**
13.   Group j cells in a cluster;
14.   Iterate step 7–13 to find the left (k-j) cells;
15.   **end**
16.   Remove chosen cells from $\{C_i\}$;
17. **end**
18. **return** cluster sets.

power consumption. Second, to evaluate the comparison with previous EDAC technique, an EDAC processor is proposed. It employs the latest razor-style circuits [5] to detect timing error and correct it by error mask method. For fair comparison, it still needs to insert clock gate in the non-critical paths. Third, we realize SGERC with the same RTL code and its implementation details are provided in Table III. The area overhead mainly comes from the EDAC circuits and short path fixing, which takes 5.81% over the baseline design. And we employ 24 clusters to replace the original 163 critical flip-flops. Compared with EDAC design, 1.45% area overhead is due to the newly added data-driven clock gating logic in the critical path.

Table III. SGERC processor implementation details

Technology Node	SMIC 40 nm LL
Voltage Range	0.6∼1.1 V
Total Number of Logic Gates	23268
Target Clock Frequency	20 MHZ @0.6 V 236 MHZ @1.1 V
Number of clusters	24
Number of replaced flip-flops	163
Detection Window	25% of System Clock
Total Core Area Overhead	5.81% over baseline Design 1.45% over EDAC Design

4.2 Simulation results

Wide-voltage operation. During the voltage scaling, the energy efficiency of three designs is evaluated. To account for PVTA variation, we add 30% design margin for the baseline design compared with the nominal operating voltage as Razor II [1]. And EDAC and SGERC designs work at the point of first failure (POFF) to get the simulation results. Meanwhile, we choose Drystone benchmark as the test case which can cover all the critical paths in CK802 processor.
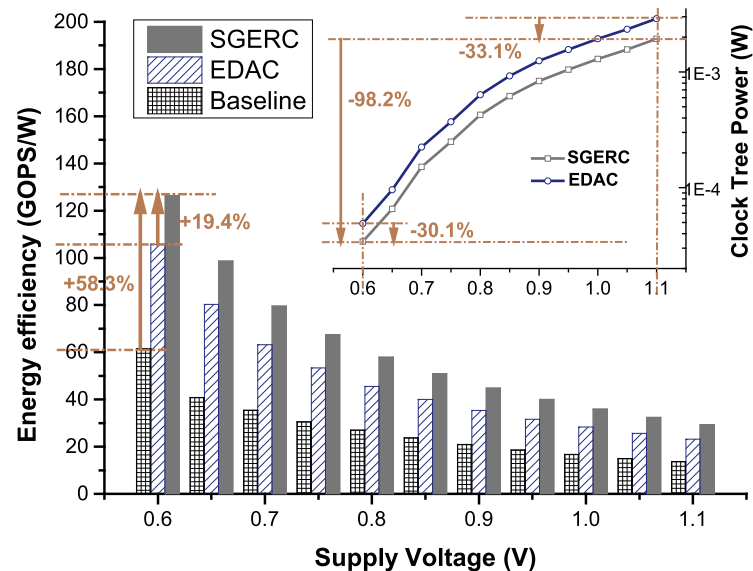

Fig. 6. Comparison of energy efficiency and clock tree power during the voltage scaling.

As shown in Fig. 6, SGERC design improves the energy efficiency by 58.3% compared with the baseline design at 0.6 V. The power benefits come from the protection of error resilient circuits because it can work at a lower voltage without design margins to achieve the same throughput. Further, its promotion over EDAC design is 19.4% by adding the clock gating and eliminating the previous error detection circuits. The sub-figure also shows that the clock tree power of SGERC is always lower than EDAC design and it can be reduced by 30.1% at 0.6 V.

Overclocking operation. By improving the working frequency and operating beyond the POFF, SGERC and EDAC designs evaluate their error count and energy efficiency. Fig. 7 shows that the working frequency increases by 33% (@37.5 ns). Meanwhile, the error count of SGERC increases and is 4.14 times as EDAC's. And its POFF is at 47 ns while EDAC's is 45.5 ns. This is because the clock gates are newly added into critical path of EDAC design and face a more tight timing constraints. However, the energy efficiency of SGERC is always larger than EDAC circuits with an improvement of nearly 19%. Because the improvement of energy efficiency mainly comes from the clock tree power reduction. Further, during the voltage scaling, its energy efficiency changes little with 0.18% and achieves a maximum value at 24.10 MHZ. Because it realizes the timing error mask technique with only one-cycle correction penalty. Finally, it achieves the conclusion that SGERC method can have larger energy efficiency during frequency scaling even though its error rate is higher.
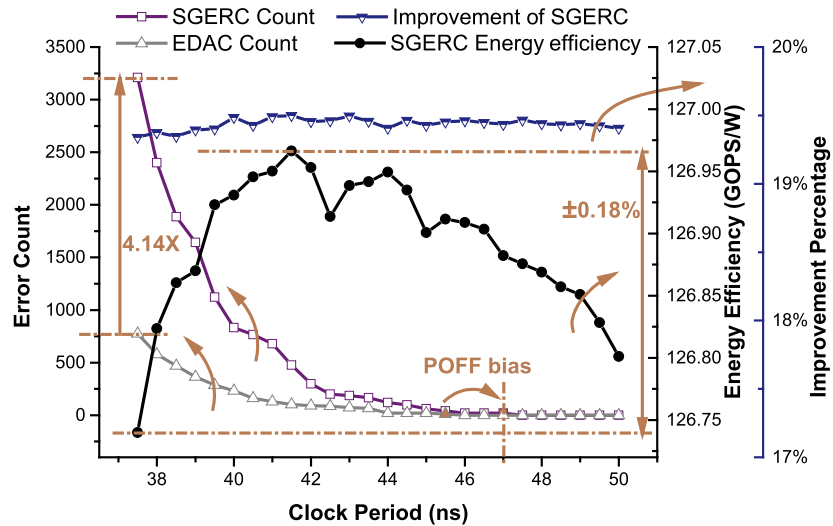
Fig. 7. Error count and energy efficiency by overclocking operation.

4.3 Comparison with other works

Since few works introduce clock gating to EDAC system, we compare our work with the latest EDAC application as listed in the Table IV. First, SGERC design adopts Flip-flop/latch to sample data, which is easier to achieve timing closure by commercial tools than latch-based design. Second, for the cluster of k cells ($k = 8$), the average number of transistors added for each cell is 13. And 12 of them are due to the XOR-gate at each latch which are shared by the clock gating mechanism. Transition detector in [11] occupies 32 extra transistors while its total area overhead is only 6.9%. This is because it employs a detection window about 5% of system clock period so that the hold fixing cost is much less. And smaller detection window means that the system is more vulnerable to PVTA variations. Third, the detection window of SGERC is 25% and its area overhead is the lowest. Finally it achieves the highest energy efficiency about 58.3% under the help of clock gating. Besides, the TESCG we propose can also be embedded into the clock network of other EDAC design.

Table IV. Comparison with previous EDAC works

Design	[11]	[12]	iRazor [5]	This paper
Cell Type	Flip-Flop	Latch	Flip-Flop/ Latch	Flip-Flop/ Latch
Processor	32 bit, 6 stage	16 bit, 5 stage	ARM Cortex-R4	C-SKY CK802
Extra Transistor	32	24	3	13
EDAC cell/ Total Cell	503/ 2976	57/ 445	1115/ 12875	FF:163/1148 SGERC:24
Area Overhead	6.9%	8.3%	11.9%	5.81%
Energy Efficiency	43%	38%	46%	58.3%

5 Conclusion

In this paper, we propose a self-gated sequential cell cluster which supports the EDAC mechanism. It can eliminate the error detection logic for every critical registers before and provide the data-driven clock gating by an error resilient clock gate called TESCG. TESCG can restore the clock signals by itself when timing violation occurs with only two transistors added. Since SGERC needn't modify processor architecture, it can be integrated automatically into the EDAC system. We implement it in CK802 processor and the simulation results show a total 58.3% improvement in energy efficiency compared with baseline design and 19.4% over EDAC design.

Acknowledgments