Abstract— This paper proposes a new technique for implementing a high speed multiplier using self evolving hold logic and razor flip flop. The overall performance of the Digital multiplier systems depends on throughput of the multiplier. The negative bias temperature instability effect occurs when a pMOS transistor is under negative bias (\(V_{gs} = -V_{dd}\)) increasing the threshold voltage of a pMOS transistor and reducing the multiplier speed. Similarly, positive bias temperature instability occurs when an nMOS transistor is under positive bias. Both effects degrade the speed of the transistor and in the long term, the system may fail due to timing violations. Therefore, it is required to design reliable high-performance multiplier. This paper, implements an aging aware multiplier design with a novel self evolving hold logic (SEHL) circuit. The multiplier is able to provide the higher throughput through the variable latency and can adjust the SEHL circuit to lesser performance degradation that is due to the aging effect. The proposed design can be applied to the column bypass multiplier and as well as row bypass multiplier.

Keywords—Self evolving hold logic (SEHL), Negative bias temperature instability (NBTI), Positive bias temperature instability (PBTI), Reliable multiplier, Variable latency.

I. INTRODUCTION

Digital multipliers are among the most critical arithmetic functional units in many applications, such as the Fourier transform, Discrete cosine transforms, and Digital filtering. The throughput of these applications depend on multipliers. If the multipliers are too slow, the performance of entire circuits will be reduced.

Furthermore, NBTI occurs when a pMOS transistor is under negative bias (\(V_{gs} = -V_{dd}\)). In this situation, the interaction between inversion layer holes and hydrogen-passivated Si atoms breaks the Si–H bond generated during the oxidation process, generating H or H2 molecules. When these molecules diffuse away, interface traps are left. The accumulated interface traps between silicon and the gate oxide interface result in increased threshold voltage (\(V_{th}\)), reducing the circuit switching speed. When the biased voltage is removed, the reverse reaction occurs, reducing the NBTI effect. However, the reverse reaction does not eliminate all the interface traps generated during the stress phase, and \(V_{th}\) is increased in the long term. Hence, it is important to design a reliable high performance multiplier.

The corresponding effect on an nMOS transistor is PBTI, which occurs when an nMOS transistor is under positive bias. Compared with the NBTI effect, the PBTI effect is much smaller on oxide/polycrystal transistors, and therefore is usually ignored. However, for high-k/metal-gate nMOS transistors with significant charge trapping, the PBTI effect can no longer be ignored. In fact, it has been shown that the PBTI effect is more significant than the NBTI effect on 32-nm high-k/metal-gate processes [4].

Traditional method to mitigate the aging effect has been already done in [8] and [9], using such things as guardbanding and gate over sizing; however, this approach can be very pessimistic and area and power inefficient. To avoid this problem, many NBTI-aware methodologies have been proposed. An NBTI aware technology mapping technique was proposed in [12] to guarantee the performance of the circuit during its lifetime. In [13] an NBTI-aware sleep transistor was designed to reduce the aging effects on pMOS sleep-transistors, and the lifetime stability of the power-gated circuits under consideration was improved. Wu and Marculescu proposed a point logic restructuring and pin reordering method, which is based on detecting functional symmetries and transistor stacking effects. They also proposed an NBTI optimization method that considered path sensitization, dynamic voltage scaling and body-basing techniques were proposed to reduce power or extend circuit life. These techniques, however, require circuit modification or do not provide optimization of specific circuits. Traditional circuits use critical path delay as the overall circuit clock cycle in order to perform correctly. However, the probability that the critical paths are activated is low. In most cases, the path delay is shorter than the critical path. For these noncritical paths, using the critical path delay as the overall cycle period will
result in significant timing waste. Hence, the Variable-latency design was proposed to reduce the timing waste of traditional circuits.

The variable-latency design divides the circuit into two parts: 1) shorter paths and 2) longer paths. Shorter paths can execute correctly in one cycle, whereas longer paths need two cycles to execute. When shorter paths are activated frequently, the average latency of variable-latency designs is better than that of traditional designs. For example, variable-latency adders were proposed using the speculation technique with error detection and recovery [5]. A short path activation function algorithm was proposed to improve the accuracy of the hold logic and to optimize the performance of the variable-latency circuit. An instruction scheduling algorithm was proposed to schedule the operations on non-uniform latency functional units and improves the performance of Very Long Instruction Word processors. A variable-latency pipelined multiplier architecture with a Booth algorithm was also proposed. In [14], process-variation tolerant architecture for arithmetic units was proposed, where the effect of process-variation is considered to increase the circuit yield. In addition, the critical paths are divided into two shorter paths that could be unequal and the clock cycle is set to the delay of the longer one. These research designs were able to reduce the timing waste of traditional circuits to improve performance, but they did not consider the aging effect and could not adjust themselves during the runtime. A variable-latency adder design that considers the aging effect was proposed in [18] and [19]. However, no variable-latency multiplier design that considers the aging effect and can adjust dynamically has been done.

This paper has been organized in the following way, we propose an aging-aware reliable multiplier design with a novel SEHL circuit. The multiplier is based on the variable-latency technique and can adjust the SEHL circuit to achieve reliable operation under the influence of NBTI and PBTI effects. To be specific, the contributions of this paper are summarized as follows:

1) Novel variable latency multiplier architecture with an SEHL circuit. The SEHL circuit can decide whether the input patterns require one or two cycles and can adjust the judging criteria to ensure that there is minimum performance degradation after considerable aging occurs;

2) Comprehensive analysis and comparison of the multiplier’s performance under different cycle periods to show the effectiveness of our proposed architecture;

3) An aging aware reliable multiplier design method that is suitable for large multipliers;

4) The experimental results shows the effectiveness of the multiplier with 64x64 input bits.

II. PRELIMINARIES

A. Column Bypassing Multiplier

A column-bypassing multiplier is an improvement on the normal array multiplier (AM). The AM is a fast parallel AM and is shown in Fig. 1. The multiplier array consists of (n−1) rows of carry save adder (CSA), in which each row contains (n − 1) full adder (FA) cells. Each FA in the CSA array has two outputs: 1) the sum bit goes down and 2) the carry bit goes to the lower left FA. The last row is a ripple adder for carry propagation.

![Fig. 1 4 × 4 normal AM.](image)

The FAs in the AM are always active regardless of input states. A low-power column-bypassing multiplier design is proposed in which the FA operations are disabled if the corresponding bit in the multiplicand is 0. Fig. 2 shows a 4x4 column-bypassing multiplier. Supposing the inputs are 1010 * 1111, it can be seen that for the FAs in the first and third diagonals, two of the three input bits are 0: the carry bit from its upper right FA and the partial product \( a_i b_j \). Therefore, the output of the adders in both diagonals is 0, and the output sum bit is simply equal to the third bit, which is the sum output of its upper FA. Hence, the FA is modified to add two tristate gates and one multiplexer.
The multiplicand bit $a_i$ can be used as the selector of the multiplexer to decide the output of the FA, and $a_i$ can also be used as the selector of the tristate gate to turn off the input path of the FA. If $a_i$ is 0, the inputs of FA are disabled, and the sum bit of the current FA is equal to the sum bit from its upper FA, thus reducing the power consumption of the multiplier. If $a_i$ is 1, the normal sum result is selected.

**B. Row-Bypassing Multiplier**

A low-power row-bypassing multiplier is also the best proposal to reduce the activity power of the AM. The operation of the low-power row-bypassing multiplier is similar to that of the low-power column-bypassing multiplier, but the selector of the multiplexers and the tristate gates use the multiplicator. Fig. 3 is a $4 \times 4$ row-bypassing multiplier. Each input is connected to an FA through a tristate gate. When the inputs are $1111 \cdot 1001$, the two inputs in the first and second rows are 0 for FAs. Because $b_1$ is 0, the multiplexers in the first row select $a_i b_0$ as the sum bit and select 0 as the carry bit. The inputs are bypassed to FAs in the second rows, and the tristate gates turn off the input paths to the FAs. Therefore, no switching activities occur in the first-row FAs; in return, power consumption is reduced. Similarly, because $b_2$ is 0, no switching activities will occur in the second-row FAs. However, the FAs must be active in the third row because the $b_3$ is not zero. But as compared with the column bypassing multiplier row bypassing multiplier require more area and power. So it is better to choose column bypassing multiplier.

**C. Variable-Latency Design**

The variable-latency design was proposed to reduce the timing waste occurring in traditional circuits that use the critical path cycle as an execution cycle period as shown in Fig. 4. The basic concept is to execute a shorter path using a shorter cycle and longer path using two cycles. Since most paths execute in a cycle period that is much smaller than the critical path delay, the variable-latency design has smaller average latency. Fig. 4 is an 8-bit variable-latency ripple carry adder (RCA). $A_8-A_1$, $B_8-B_1$ is 8-bit inputs, and $S_8-S_1$ are the outputs. Supposing the delay for each full adder is one, and the maximum delay for the adder is 8. Through simulation, it can be determined that the possibility of the carry propagation delay being longer than 5 is low. Hence, the cycle period is set to 5, and hold logic is added to notify the system whether the adder can complete the operation within a cycle period.
III. PROPOSED RELIABLE MULTIPLIER

A. Proposed model

The multiplier architecture, which includes two m-bit inputs (m is a positive number), one 2m-bit output, one column-bypassing multiplier, 2m 1-bit Razor flip-flops and an SEHL circuit as shown in Fig.5.

![Multiplier design with Self evolving hold logic](image)

The column-bypassing multiplier can be examined by the number of zeros in the multiplicand to predict whether the operation requires one cycle or two cycles to complete. When input patterns are random, the number of ones in the multiplicand follows a normal distribution. According to the bypassing selection, the input signal of the SEHL in the architecture with the column-bypassing multiplier is the multiplicand.

B. Razor Flip Flop

Razor flip-flops can be used to detect whether timing violations occur before the next input pattern arrives as shown in Fig.6. A 1-bit Razor flip-flop contains a main flip-flop, shadow latch, XOR gate, and multiplexer. The main flip-flop catches the execution result for the combination circuit using a normal clock signal, and the shadow latch catches the execution result using a delayed clock signal, which is slower than the normal clock signal. If the latched bit of the shadow latch is different from that of the main flip-flop, this means the path delay of the current operation exceeds the cycle period, and the main flip-flop catches an incorrect result. If errors occur, the Razor flip-flop will set the error signal to 1 to notify the system to reexecute the operation and notify the SEHL circuit that an error has occurred.

![Razor Flip Flop](image)

C. Self Evolving Hold Logic

The SEHL circuit is the key component in variable-latency multiplier. Fig.7 shows the details of the SEHL circuit. The SEHL circuit contains an aging indicator, two judging blocks, one multiplexer, and one D flip-flop. The aging indicator indicates whether the circuit has suffered significant performance degradation due to the aging effect. When input patterns arrive, the column-bypassing multiplier, and the SEHL circuit execute simultaneously.

![Self Evolving Hold Logic](image)

According to the number of zeros in the multiplicand, the SEHL circuit decides if the input patterns require one or two cycles. If the input pattern requires two cycles to complete, the SEHL will output 0 to disable the clock signal of the flip-flops. Otherwise, the SEHL will output 1 for normal operations. When the column-bypassing multiplier finishes the operation, the result will be passed to the Razor flip-flops.
IV. SIMULATION RESULTS

The below figure shows the Block diagram of the aging-aware multiplier.

![Block diagram of the aging-aware multiplier](image)

Fig. 8 Block diagram of the aging-aware multiplier

The below figure shows the RTL Schematic of the 64 x 64 aging-aware multiplier.

![RTL Schematic of the aging-aware multiplier](image)

Fig. 9 RTL Schematic of the aging-aware multiplier

The below figure shows the Technology Schematic of the 64 x 64 aging-aware multiplier.

![Technology Schematic of the aging-aware multiplier](image)

Fig. 10 Technology Schematic of the aging-aware multiplier.

The below Fig. 11 shows the Simulation output waveform of the 64 x 64 aging-aware multiplier when the multiplier needs only one clock cycle and Fig. 12 shows the simulation output waveform of the 64 x 64 aging-aware multiplier when the multiplier needs two clock cycles.

![Simulation output waveform of the aging-aware multiplier](image)

Fig. 11 Simulation output waveform of the aging-aware multiplier when it takes one clock cycle.
V. CONCLUSION

This paper proposed an aging-aware variable-latency multiplier design with the SEHL. The multiplier is able to adjust the SEHL to mitigate performance degradation due to increased delay. The proposed variable latency multipliers have less performance degradation because variable latency multipliers have less timing waste, but traditional multipliers need to consider the degradation caused by the BTI effect. The experimental results show that the proposed architecture with 64 x 64 multiplication with CLA as last stage instead of Normal RCA adder to reduce the delay to some more extent. It will decrease the delay and improve the performance compared with previous designs.

REFERENCES


