

MOSVLSI DESIGN
TERM PROJECT 8-BIT WALLACE TREE MULTIPLIER
SUBMITTED BY: JHALAK PATEL

CONTENTS

1. OVERVIEW.....	2
2. DESCRIPTION.....	2
3. DESIGN COMPONENTS.....	3-4
3.1 Mirror Adder	
3.2 True Single Phase Clocked Flip Flop	
3.3 Double Clocked CMOS Latch	
4. DESIGN BLOCKS.....	4-6
4.1 Wallace Tree	
4.2 Latency Predictor	
4.3 Merge Adder	
4.4 Clock Tree	
5. ANALYSIS AND CONCLUSION.....	7
5.1 Area Analysis	
5.2 Power Analysis	
5.3 Timing Analysis	

FIGURES

1. Wallace Tree Multiplier with Adaptive Clocking.....	3
2. 8x8 Wallace Tree Multiplier.....	4
3. Latency Predictor Block.....	5
4. Clock Tree for the Complete Design.....	6
5. Floor Planning.....	7

1. OVERVIEW

In this project, we have implemented 8bit Wallace tree multiplier at scaled supply voltage with adaptive clock stretching. Multipliers are one of the most important parts in the signal processing or other computationally intensive applications. Therefore, we aimed at designing multiplier which operates at high speed, low power and occupies optimal area. We have used 180nm technology for the project. For this project, we have created the library of blocks from scratch. We have used various novel blocks to realize the completed designs such as Mirror Adder, True Single Phase Clock Flip Flop, Doubled CMOS Clocked latch. We have done the characterization of the component, the schematic and layout implemented using Cadence tools. For verifying the HSPICE Netlist, we have used NanoSim for functional simulation as well as for Power, Area and timing analysis.

Process Technology	180nm Technology
Final Area of Chip	162x185um-square
Number of PMOS	1358
Number of NMOS	1420
Total Number of Transistors	2778
Power Dissipation*	2.75mW
Voltage	1.8V
Clock Frequency*	333MHz

*All the Power and Timing Calculations are for Extracted Design

2. DESCRIPTION

Our design takes two 8bit binary numbers as inputs for multiplication operation. The inputs to the system are fed with a clock pulse of 3ns for the optimal operation of the design. Firstly, the inputs to the design are latched using Input Flip Flop stage which are clocked by gated clock rising edge (i.e. Coming from the Latency Predictor Block). The clocking is done through the clock coming out of clock tree which ensures that the clock is not skewed or distorted. Latched outputs of the first latch are fed to AND array to generate 64bit partial products. The partial products produced are fed to the Wallace tree which basically has Layers of Adder which compresses the inputs and propagate the carry to the next stage. The output of the Wallace tree are latched using the intermediate stage Flip Flop stage. After Sampling the output at the intermediate stage, the middle bits are decoded using the *Latency Predictor Block (LPB)*, which basically generates the enable signal to stall clock. The output of the LPB is gated with the input clock to generate the *enabled_clock* which is fed to all the three latches i.e. input, intermediate and output latch. The outputs sampled from the middle latch is fed to the merge adder. We have implemented merge adder using ripple carry adder which produces the final 16bit output which are latched by the output Flip Flop to generate the final output.

In our implementation we are implementing each multiplication with latency of 3 or 4 (with the stretched clock) clock cycles. But due to pipelining of computational block (combination logic), we are able to get higher throughput i.e. One output per clock cycle.

Advantages of Adaptive Clocking:

In our design, we have reduced the critical path in the merge adder stage, so that we can operate the clock at higher frequency at lower voltage. We also made observation that in the absence of the adaptive clocking, in order to operate the circuit at same frequency we need to apply higher supply voltage. The adaptive clocking helped us to scale down the operating voltage.

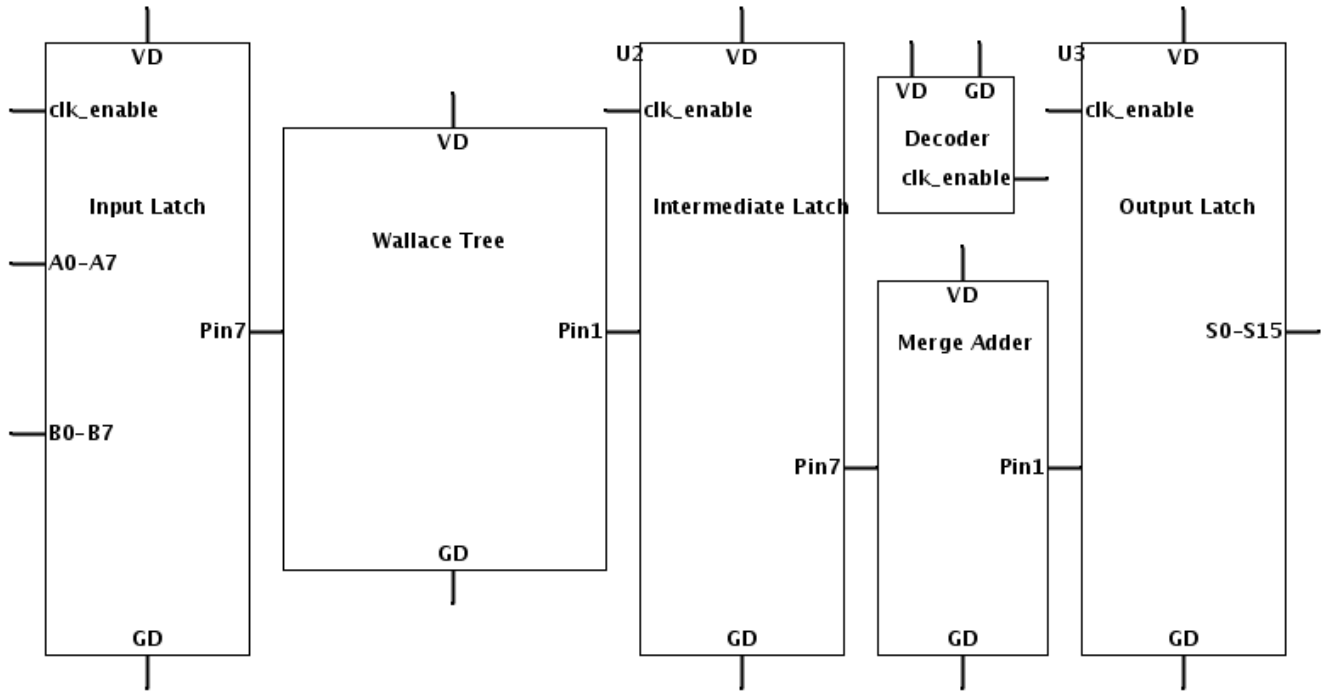


Figure1: Wallace Tree Multiplier with Adaptive Clocking (Block Diagram)

3. DESIGN COMPONENTS:

3.1 Mirror Adder

We have implemented Full adder as a Mirror Adder. Mirror adder is comparatively more faster than the normal full adder circuit. Due to its symmetric nature, rise time and fall time for the mirror is almost the same as we have properly sized our NMOS and PMOS.

Mirror Adder provides its benefits in the carry generation circuitry as there are maximum of two transistors in series for the carry generation path. We have optimized the transistors in the carry stage for optimal operation of the adder.

Due to symmetric nature of the circuit, the PMOS pullup network was designed and then mirrored across the horizontal axis to facilitate the creation of NMOS pulldown network. Which merely required changing the diffusion types and re-sizing the active areas to ensure the proper size ratio.

3.2. TSPC FlipFlop

We have implemented true single phased clock FF, which acts as a Master Slave FF. We have proposed the use of dynamic circuit based FF which comparatively has less delay than the Static CMOS based D-FlipFlop. The TSPC FF used as Edge Triggered D flip-flop operates at high speed using dynamic logic. The digital output is stored on the parasitic device capacitance while the device is not transitioning. The charge based storage limits the usage of such FF for static or low clock speed circuits. However, we were encouraged to use the dynamic circuit based FF as we planned to operate our circuit at higher clock frequency, such that the FF is clocked fast enough. As by operating at higher frequency we avoided the leakage condition for the parasitic capacitance which could have caused the FF to enter invalid states.

When the clock is low, the master stage is transparent and latches the outputs when the level goes high, the Slave stage reflects the latched output in the previous stage. For error free operation of the FF, clock pulse width should be more than delay of the slave stage.

3.3 Double C2MOS

We have implemented negative latch using Doubled C2MOS. Using Dynamic logic based latch enabled us to use it for faster switching operations. In addition to the advantage of the higher operations speed, it has only 6 transistor count, which helped in area reduction as compared to traditional D-latch formed using Static CMOS logic (i.e. NOR based or NAND based Latch)

4. DESIGN BLOCKS

4.1 Wallace Tree

Wallace tree is efficient hardware implementation for N-bit multiplication. In our case we have designed Wallace tree multiplier for two 8-bit inputs. First we have generated 64-bit partial products by ANDing the combination of input bits. In the next stage the tree reduces the partial products by compression (addition) using the 4 layers of Full Adder and Half Adders. Full Adder can be thought of as a 3:2 bit compressor which reduces the partial product inputs. At the end of the Wallace Tree stage, we get the 26-bit compressed output which can be used for further computation.

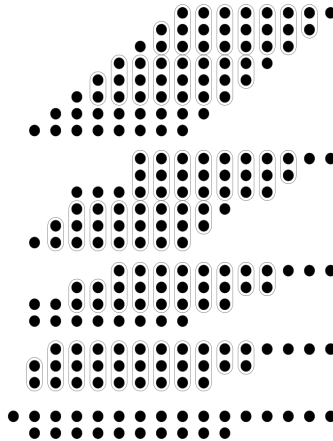


Figure 2: 8x8 Wallace Tree Multiplier (Logical Overview)

4.2 Latency Predictor

The Latency Predictor block decodes the intermediate input to predict if the current computation requires clock stretching or not. We have XORed the intermediate input to evaluate the carry propagation bit, which is ANDed with the corresponding output of the other XOR gates. We have used positive edge triggered D-Flop Flop to feedback the enable output bit of the previous cycle to generate the correct value of enable in the current cycle. It also ensures that for the next clock cycle, enable bit is set to '1' again, so that we can get accurate stalling for one additional clock cycle.

The Latency Predictor block also includes a negative D-latch (C2MOS latch) which latches the enable bit output at the negative clock level, thus it remains latched during the next positive clock cycle.

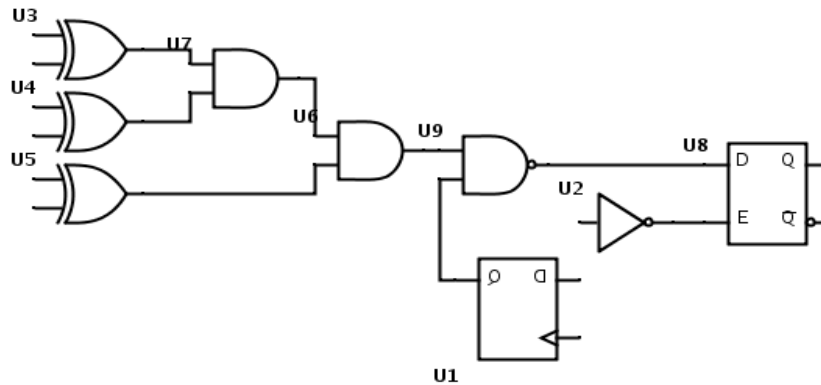


Figure 3: Latency Predictor Block

The circuit allows computation of the enable signal before the next set of input arrives i.e. At the next rising edge of the clock, which is extremely critical for the success of our design methodology. The value of the enable latched in the negative clock cycle is used to determine whether the output register will be written at the next rising edge of the clock or be delayed by one clock cycle to implement clock stretching. We have used clock gating to generate the final enabled clock signal to disable the write operation to the input, output and intermediate latch.

4.3 Merge Adder

We have used Ripple Carry adder circuitry for adding the latched output bit of the Wallace tree. It acts as a merge adder which adds the 25-bit numbers. Each full adder inputs a C_{in} which is C_{out} of the previous adder such that each carry bit "ripples" to the next full adder. We have first adder as a Half Adder as C_{in} to the first stage is '0'.

We have chosen ripple carry adder as it was simple which allowed us fast design time. Also the transistor count in case of ripple carry adder is less as compared to other adders such as carry-select adder.

Delay Analysis for the Merge Adder

Merge Adder has 25 Adder Stages which need to propagate the carry to generate the final stable outputs which can be latched at the later stage. In order to evaluate the worst case

time(criticaldelay)forthemergeadderstage,consideringthecarryispropagatedthroughalltheadderstages:

$$t_{delay} \Rightarrow 2(\text{from the input to carry in the first adder}) + 19 \times 3 (\text{for carry propagation in the later adders}) \Rightarrow \mathbf{57 \text{ gated delays}}$$

In order to minimize the worst case delay or the critical path delay, we have implemented a Latency Predictor Block which basically analyzes middle bits (C8-P9 C9-P10 C10-P11 C11-P12) for the carry propagation, we stall the clock for an extra clock cycle for the Merge Adder computation if there is carry propagation. Thus the total delay remains 57 gated delays, which takes two clock cycles delay for calculations.

In case of no carry propagation, we consider the critical path up-to the middle adders such that the total delay is 28 gated delay, which enables us to operate the circuit at a higher clock rate.

Clock Period Calculations

$$\text{Probability of Carry Propagation} \times 2 \text{ clock cycle} + \text{Probability of No Carry Propagation} \times 1 \text{ clock cycle} \Rightarrow 0.125 \times 2 + 0.875 \times 1 \Rightarrow \mathbf{1.125 \text{ CLOCK (EFFECTIVE CLOCK CYCLE)}}$$

4.4 Clock Tree

The clock enable signal (*clk_enable*) generated after the clock gating is used to synchronize all the three latches used in the design. The *clk_enable* signal is fed to optimal size buffer stages to generate the final clock outputs for all the individual latches. The buffer stages are laid down in the layout such that there is no distortion in clock strength of the clock driving different latches. Our layout strategy also ensures minimum skew between the different clock driving the latches.

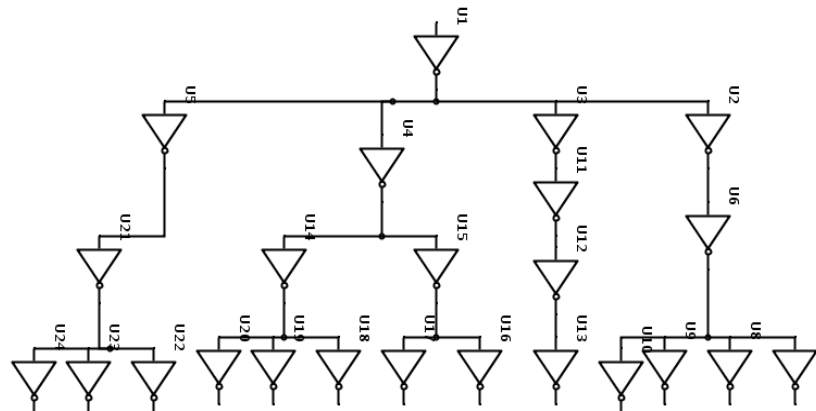


Figure 4: Clock Tree for the Complete Design

5. ANALYSIS AND CONCLUSION:

5.1 Area Analysis

Dimension of Full Adder: 14.75 x 6.5 μm -square

Dimension of TSPCLatch: 7.5 x 7.0 μm -square

Dimension of Complete Design: 162 x 185 μm -square

Layout Approach: In the design, we have tried to optimize the performance of the circuit by laying out the all the three latches more closely to each other. We have accomplished it by laying out the Adder Tree (compressor) in the form of a 'C' shaped structure. By using such layout technique, we succeeded to place the all the three latches more closely to each other. Routing and Placement of the clock Tree also played a critical role in overall performance of the design. Due to the congestion of signals, we have used up to 4 metal layers while routing different components.

Power and Ground Routing: We have used metal 1 layer for complete Power and Ground routing. We have followed 'Grid' topology to route these signals.

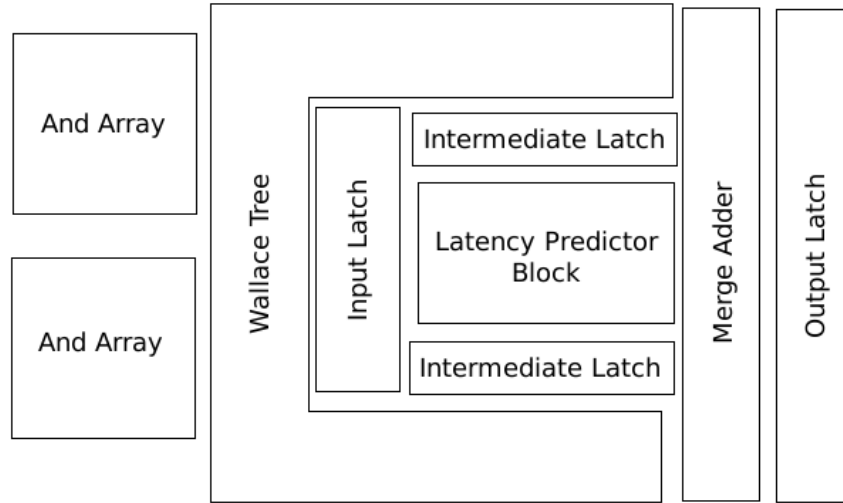


Figure 5: Floor Planning

5.2 Power Analysis

TSPC Master Slave FF Power Consumption

Average Block Power: 7.05 μ W

RMS Block Power: 65.75 μ W

Mirror Adder Power Consumption

Average Block Power: 8.83 μ W

RMS Block Power: 76.61 μ W

Complete Design Power Consumption

Average Block Power: 2.750 mW

RMS Block Power: 3.755 mW

5.3 Timing Analysis

Schematic Operating Frequency: 500 MHz

Extracted Layout Operating Frequency: 333 MHz

Extracted layout operates at a lower frequency as compared to the schematic as we have accounted for the parasitic capacitance between various metal layers and components. The extracted layout model also considers the routing delay and the resistance between various 'via's' while transitioning from one metal to another. It also accounts for the delay in supply nets i.e. Power and Ground signals due to routing topology.