Logo
Far East Journal of Electronics and Communications

http://dx.doi.org/.../...
Volume …, Number …, 2016, Pages … ISSN: 0973-6999

PIPELINED AND CLOCK GATED MAC ARCHITECTURE DESIGN AND IMPLEMENTATION

Amarveer Kaur Dhindsa[1], Rajkumar Sarma[2]

[1]amardhindsa6@gmail.com, [2]rajkumar.sarma86@gmail.com

VLSI Design, School of Electronics & Electrical Engineering, Lovely Professional University, Punjab, India

ABSTRACT

Multiply-Accumulate Unit is inevitable part of digital signal processors performing high speed operations like filtering, convolutions and so on. The optimization of the MAC for not only low power and high speed is necessary but also the feasibility of the design in terms of real time applications. The present work focuses on both the aspects where various design architectures for individual blocks are optimized for low power using gating and high speed as well as the entire MAC is synchronized to perform the complete operation in single clock cycle utilizing pipelined structure to enhance the overall speed. The individual blocks take the data on the positive edges of the clock. The design is implemented using Cadence Virtuoso as well as Cadence NCSim in the 90nm technology. The two design approaches, analog and digital respectively, are compared to show the 6 times power reduction in analog design environment as compared to its digital counterpart.

1. INTRODUCTION

Multiply-Accumulate (MAC) unit is the core of all high performance digital signal processors used in applications like speech processing, audio signal processing and transmission in digital mobile phones, seismic data processing etc. For real time signal processing, MAC unit having high speed

Received: ….; Revised: ….; Accepted: ….

Keywords: MAC, RTL, low power, Virtuoso, NCSim, pipelined.

Communicated by Rajkumar Sarma

and high throughput is desired to achieve high performance DSP system [1]. To optimize the design constraints like speed and power for improved performance, the constituent blocks of MAC like multiplier, adder etc. has to be amended. The optimization can be done at the block level or at the complete unit level. The work done so far in this field has been limited to digital approach which optimizes the designs at entire unit level.

The idea of the present research work is to design pipelined and block enabled MAC architecture in analog design environment (Virtuoso). The optimization achieved compared to the digital design is noteworthy as each individual block is handcrafted at the transistor level. A gated MAC unit increases the overall power efficiency of the designs the limited energy of the battery operated portable products puts the constraint on the power consumption. The pipelined architecture increases the speed of the MAC operation by fetching the data in advance. The synchronous buffered pipeline architecture has been used to perform the entire MAC operation in single clock cycle.

The blocks of the MAC has been modified to reduce the transistor count and hence to optimize the layout area, which in turn reduce parasitic and contribute to lower power consumption and higher speed operation.

2. DESIGN APPROACH OF THE PROPOSED MAC UNIT

A conventional MAC unit consists of multiplier and accumulator that contain the sum of the previous consecutive products. The basic operation of MAC is to compute the product of two numbers x_i and y_i and add the result to the previously stored result from the last multiplication as given below:

$$F = \sum_{i=0}^{n-1} x_i y_i$$

where i denotes the range of the values.

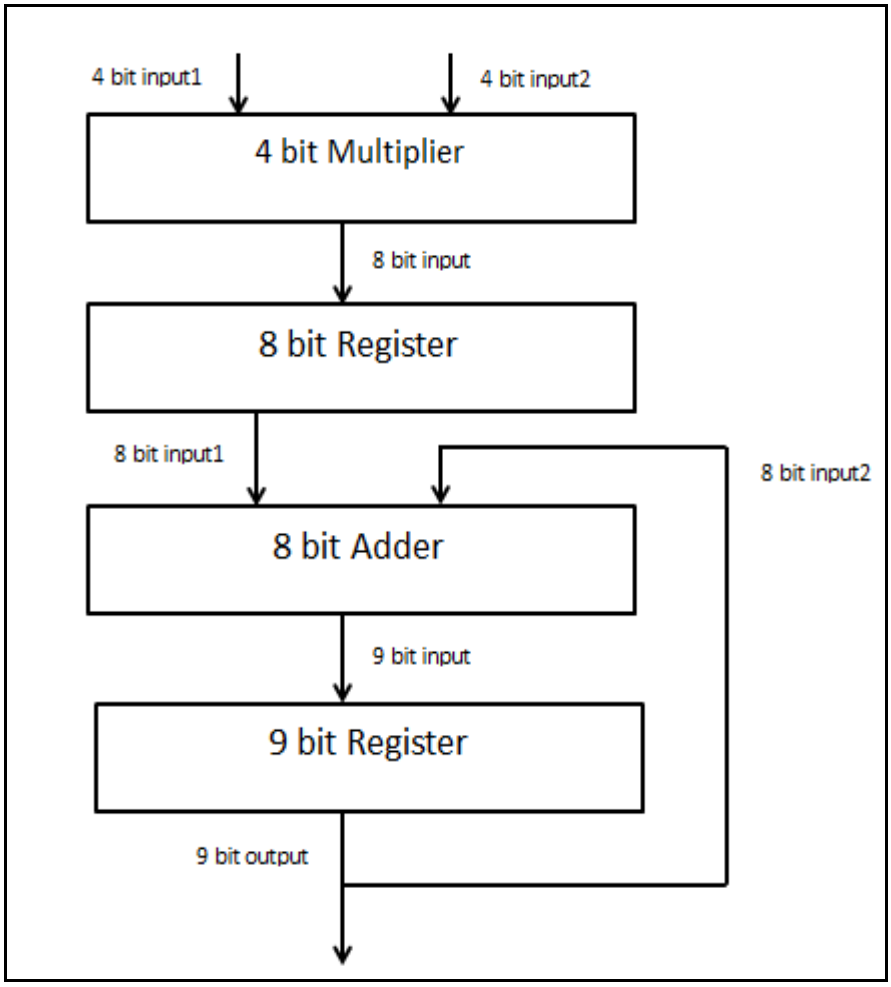

Figure 1: 4 bit MAC Architecture

The MAC inputs are fetched from the memory location and fed to multiplier block of the MAC, which will perform multiplication and give the result to adder which will accumulate the result and then will store the result into a memory location (register). This entire process is to be achieved in a single

clock cycle [2].The performance of the MAC can be increased by optimizing the basic blocks that are used in the design of the overall MAC unit [3].

2.1 Multiplier

In most of the signal processing algorithms, multiplication is a fundamental operation. Multipliers have large area, long latency and consume considerable power. System performance is generally determined by the performance of the multiplier because the multiplier is generally the slowest element in the system. Hence, optimizing the speed of the multiplier is a major concern for our design.

Australian Computer Scientist Chris Wallace suggested a fast technique to execute multiplication in 1964 [4]. The amount of hardware needed to perform this style of multiplication is large but the delay is near ideal.

Every bit of one number is multiplied with every bit of another using the AND gates. These partial products are then used as input for the adders connected in the Wallace tree architecture. The partial products of equal weights are grouped together to form a single stage. The sums of each stage are kept in the same stage and the carry is propagated to the next stage as its weight increases by 1.

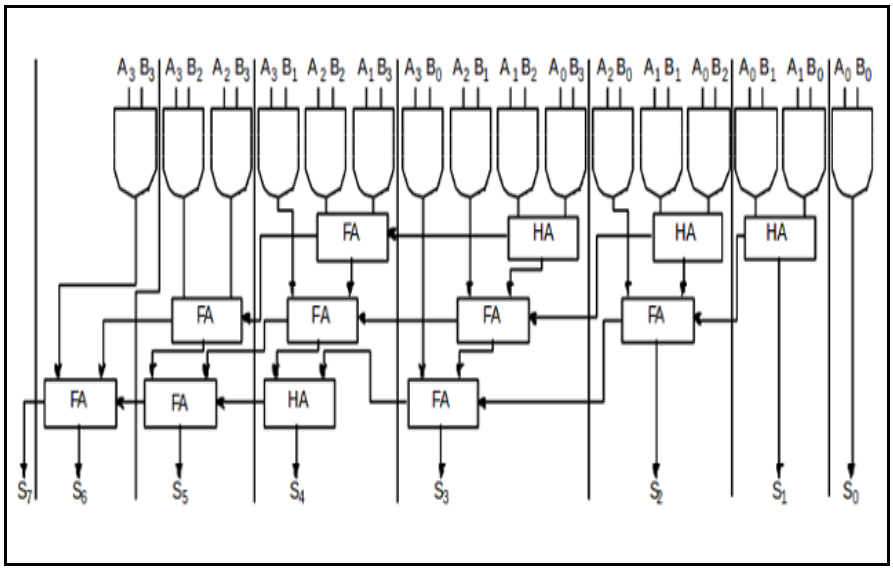

Figure 2: 4-bit Wallace Tree Multiplier

The Wallace tree multiplier is chosen because of the minimum PDP and a slight increase in the no. of transistors as compared to the advantages. The comparison is given in the table 1 given below [5].

Table 1: Comparison of Multiplier designs

Multiplier Type	Logic Style	Delay (ns)	Power (µW)	PDP (fJ)	No. of Transistor
Array	CMOS	8.300	10.73	89.06	384
	CPL	4.337	24.70	131.82	368
	DPL	4.667	19.72	92.03	448
Tree	CMOS	4.247	10.68	45.35	384
	CPL	4.105	23.61	125.25	368
	DPL	4.526	19.87	89.93	448

2.2 Accumulator

An accumulator is a register in which intermediate logic and arithmetic results are stored. If there is no register like an accumulator, it would be indispensable to write the result of each calculation (addition, multiplication, shift, etc.) to main memory, possibly only to be read right back again for use in the next operation because mathematical operations often take place in a stepwise manner, using the results from one operation as the input to the next. As the technology used for the large main memory is slower (but cheaper) than that used for a register access to main memory is slower than access to a register like the accumulator.

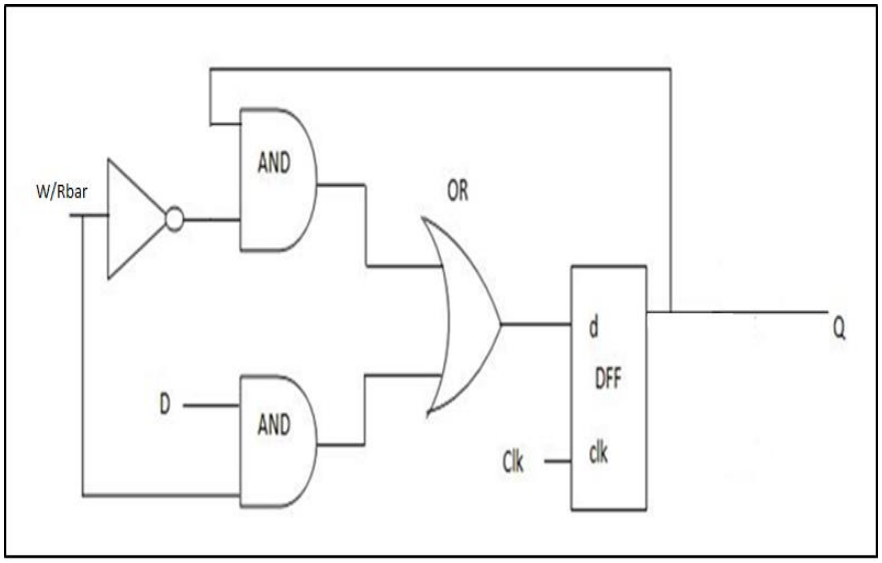

Figure 3: 1-bit register cell

The basic element constituting an accumulator or a register is a D-flip flop which can store a single bit of information. Two AND gates with clock input are also used. Hence, the register cell has three inputs write select, read select and D. The output is Q. The figure below shows the single bit register

[6]. The flip-flop is used to store the data to be used for the next cycle of operations or until the block is enabled. The D flip-flop used in this design is the positive edge triggered hybrid node flip-flop [7]. The design results in higher switching speed, low leakage currents and less power dissipation.

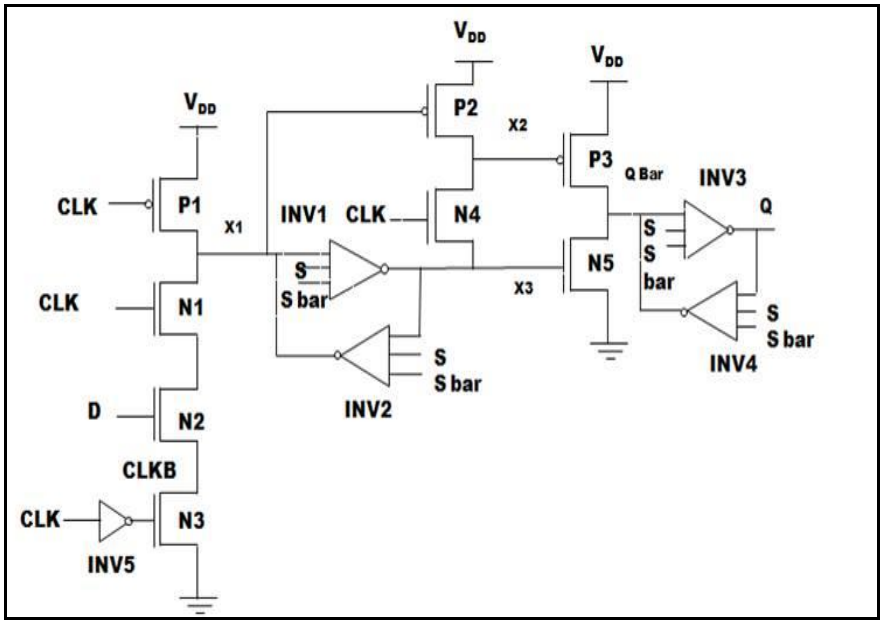

Figure 4: Positive edge triggered hybrid D flip-flop

To implement this flip-flop sleepy stack inverters are used. Sleepy stack technique is used to reduce the leakage power while increasing the speed as the states are preserved. This technique utilizes both the force stack technique as well as sleep transistor technique with high V_{th} transistors [8].

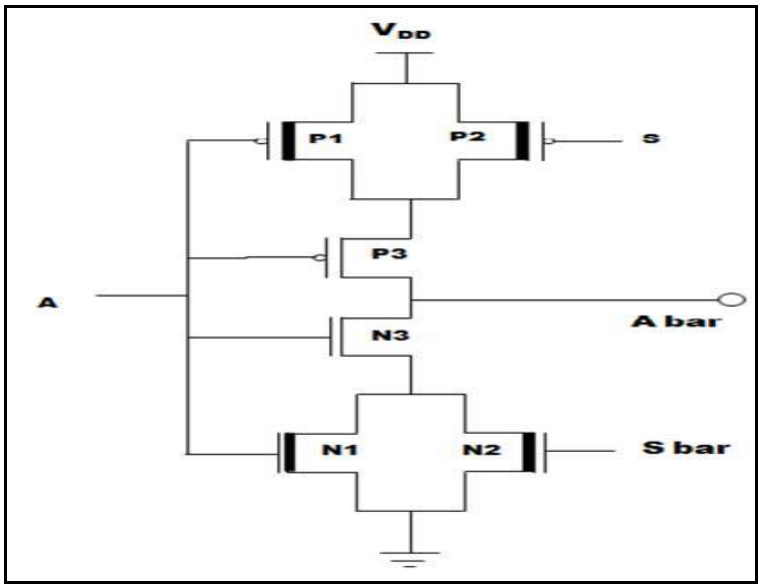

Figure 5: Sleepy Stack Inverter

The inverter works in two modes:

- **Active Mode:** In this mode the sleep signal S = 0, both the sleep transistors are turned on and the logic circuit works as a standard inverter.
- **Sleep Mode:** Here the sleep signal S = 1, the sleep transistors are turned off and the logic circuit preserves its state.

2.3 Adder

An adder also known as summer is a logic circuit which adds two numbers.

- **Half Adder:** It adds two 1-bit binary numbers and the output is Sum and Carry values. For 2-bit Wallace tree multiplier implementation and further for 8-bit multiplier half adder will be used. Also to implement 5 bit adder first block is of half adder in our design.

- **Full Adder:** It performs addition operation on 3 1-bit numbers. It takes into account the carry input also. Most of the n-bit adder architectures utilize full adders. For e.g. 5 bit ripple carry adder is used to validate the circuit performance for 2 bit MAC.

However for implementation of adder for the present research a hybrid full adder design will be utilized. In this design both CMOS logic and transmission gate logic is used.

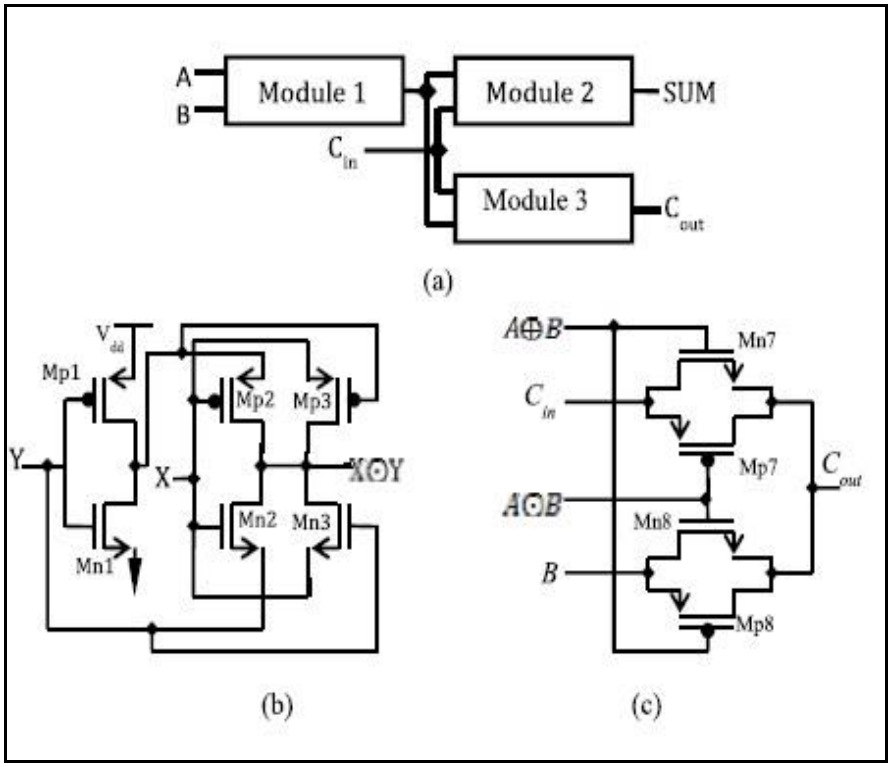

Figure 6: Hybrid Full Adder Architecture

This design is aimed at the optimization of both power and delay, minimizing the power delay product (PDP). The PDP of this design is only 0.931fJ for 1-bit adder as compared 10T FA PDP of 1.902 fJ [9].

2.4 Gating

As the feature size is scaled down, low power is the most critical issue in today's VLSI design. Gating is one of the most elegant and classic technique for reduction of dynamic power, a major contributor in total power consumption of any VLSI circuit.

$$P_{dynamic} = P_{internal} + V_{dd}^2 . f_{clk} . \alpha . c_l$$

where V_{dd} is the supply voltage, f_{clk} is the clock frequency, α represents the switching activities at nodes and c_l represents load capacitances [10].

Gating technique facilitates saving of electrical power used by digital signal processors by reducing the switching activity α. It ensures power saving by turning on a functional logic block only when required. For this first, we have to find out the delay for each stage in gating technique. Every block gets enabled only after the calculated delay. The successive blocks are disabled until the inputs are available thus saving power.

2.5 Pipelined Architecture

The basic idea for the pipelined architecture comes from everyday life. For example, water pipe continuously sends water without waiting for the water previously sent to be out. This leads to reduction in critical paths. In DSPs pipelining either reduces the power consumption at the same speed or increases the clock speed. In the buffered and synchronous pipelined architectures, "pipeline registers" are interleaved in-between functional stages, and are clocked synchronously. The time between each clock signal is set such that when the registers are clocked, the data written to them is the final result of the preceding stage [6].

The functional block in the below figure can be multiplier or adder.

Register need not to use this architecture as the register is already positive edge synchronized circuit. Whenever the positive edge of the clock is detected by the Positive Edge Detector Block the pipeline buffer will get enabled and the input available at the functional block will be passed to the output. In all other cases the functional block input will not be allowed to pass; only the previous output will be latched.

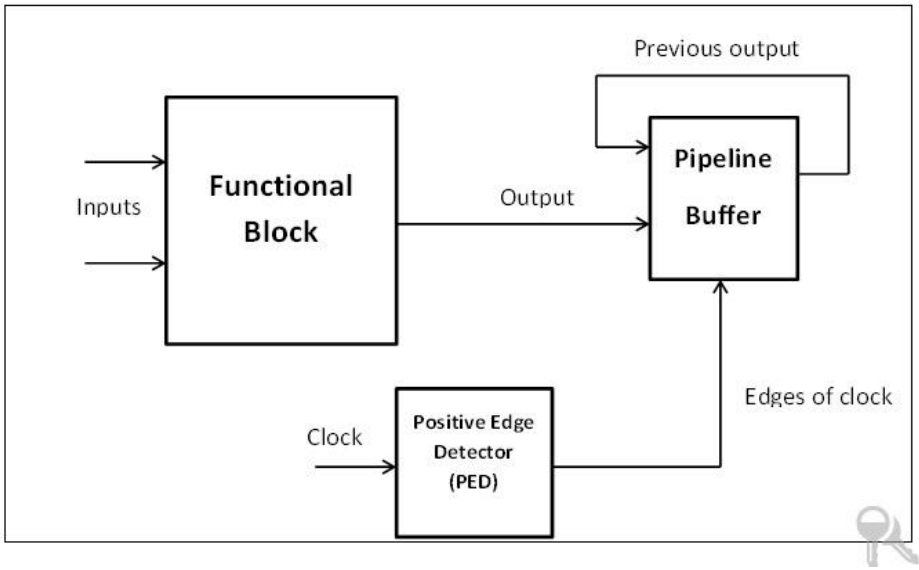

Figure 7: Pipelined and Clock Gated Architecture

3. EXPERIMENTAL RESULTS AND COMPARISONS

The 2-bit MAC unit and 4-bit MAC unit is implemented in Cadence Virtuoso using 90nm technology. The circuit is supplied with 1.1V and the operating frequency is 111 MHz. For 2 bit MAC unit, the two 2-bit numbers in binary are given as input to the multiplier and the final output is received as the 5 bit binary data from the final register block.

The two 4 bit binary numbers are input to the multiplier block of the

MAC and the final output is taken from the 9 bit register block. The intermediate signals are also plotted to verify the logic correctness and operation of the entire unit.

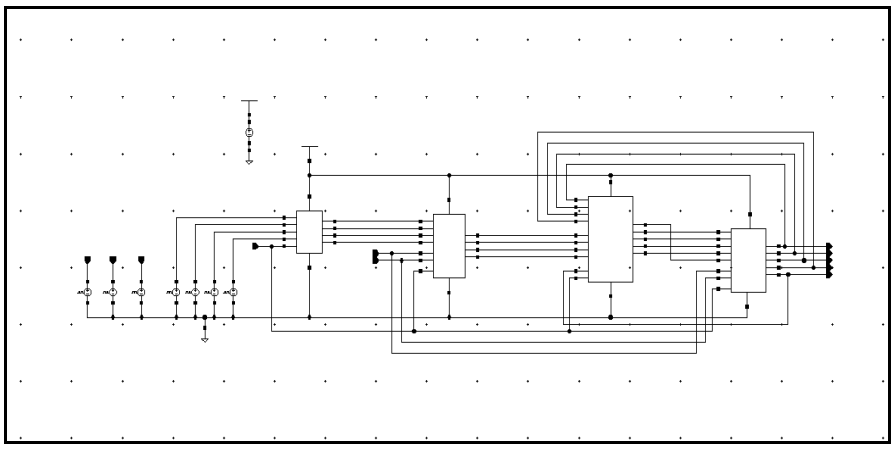

Figure 8: 2 bit MAC implementation in Virtuoso (90nm)

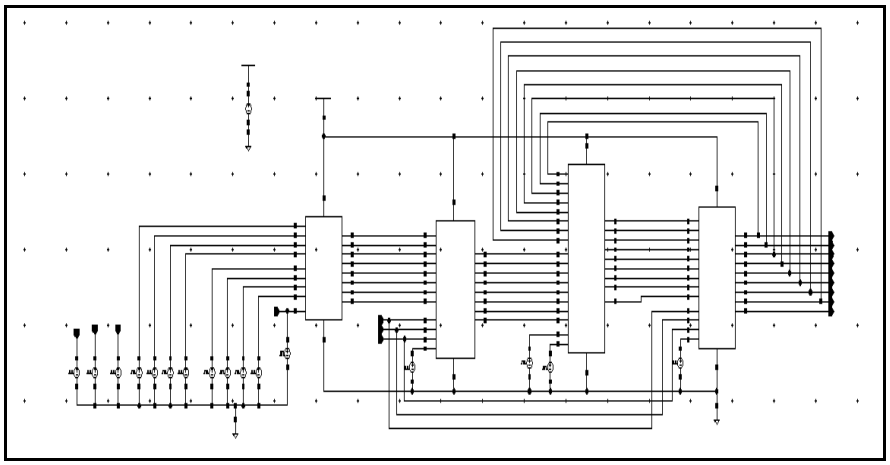

Figure 9: 4 bit MAC implementation in Virtuoso (90nm)

The simulations are run for 25ns and the power and delay are compared with the digital counterpart.

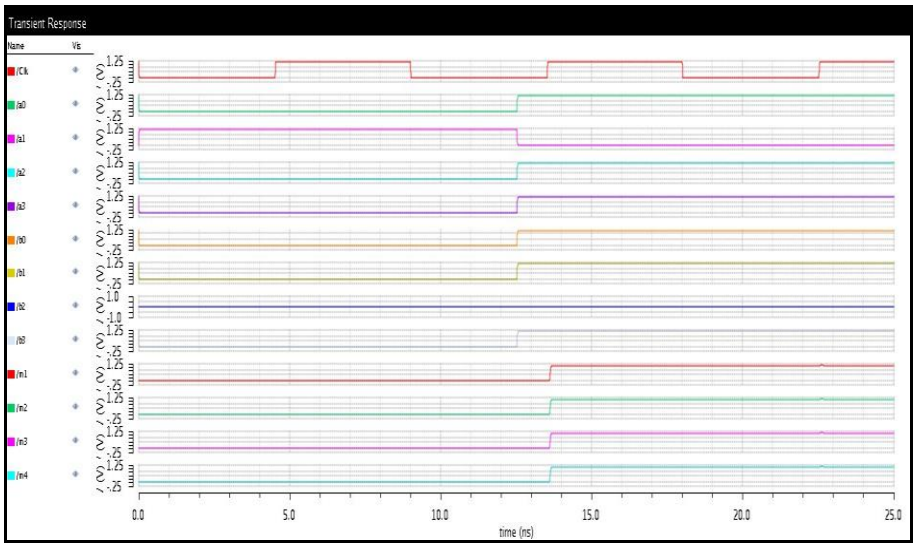

Figure 10: Simulation results for 4 bit MAC

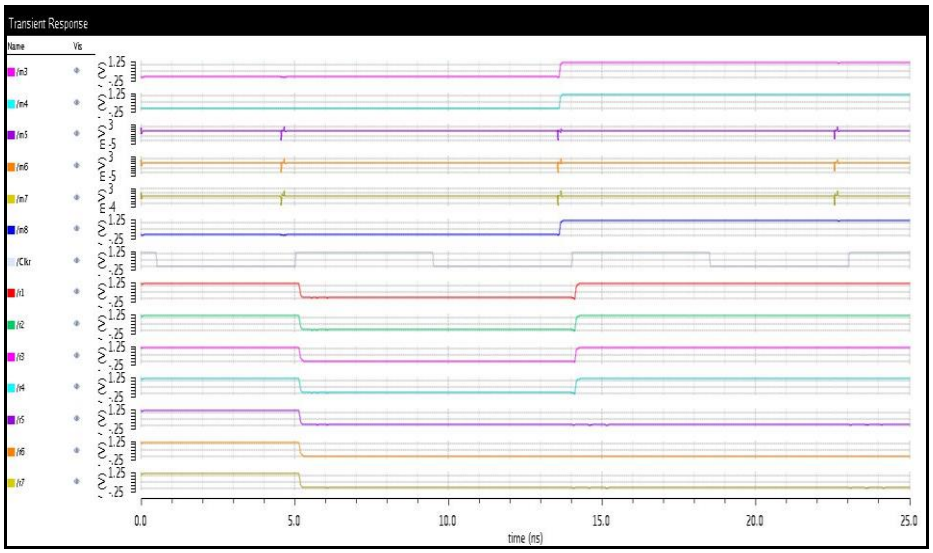

Figure 11: Simulation results for 4 bit MAC

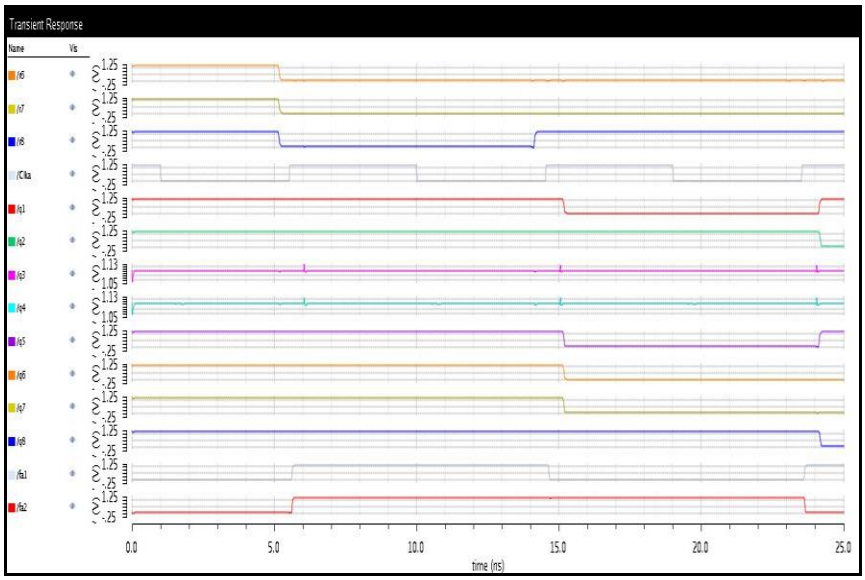

Figure 12: Simulation results for 4 bit MAC

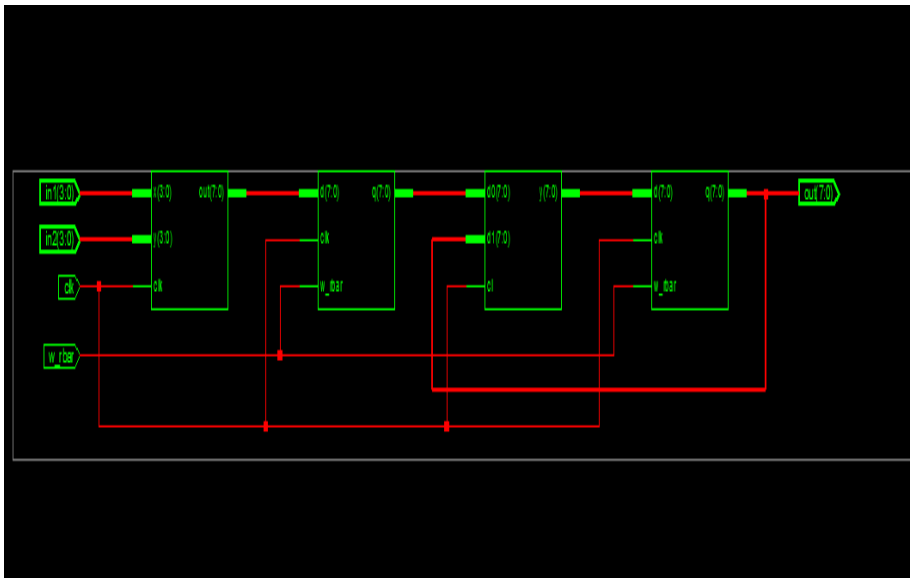

Figure 13: RTL View generated for 4 bit MAC in NCSim

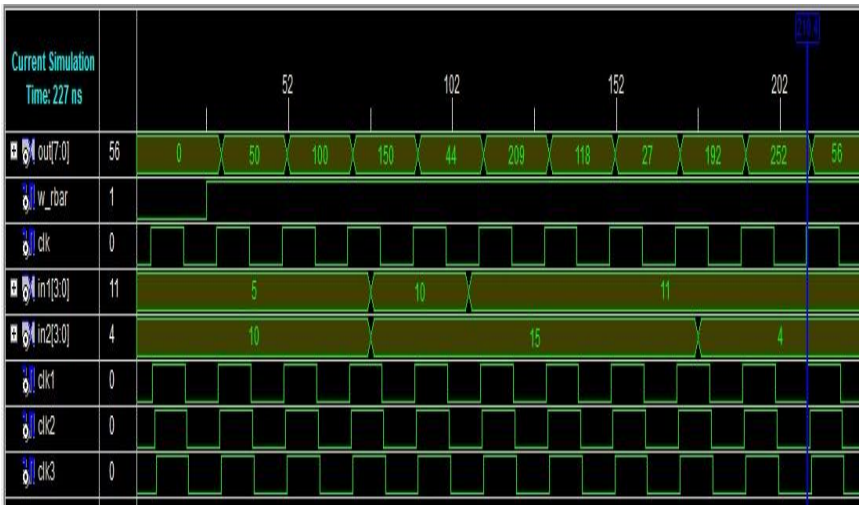

Figure 14: Simulation Result for 4 bit MAC in NCSim

The whole operation is synchronized with the rising edges of the clock signal. The digital design implementation for 2-bit MAC unit is done in the Cadence NCSim (90nm). The comparison between the two design approaches is in the table 2 given below.

Table 2: Comparison of Analog and Digital Design Implementation of 2 bit MAC

Design	**Supply Voltage**	**Total Power (µW)**
Analog Design Implementation (Virtuoso)	0.9V	9.0015
	1.1V	73.35
Digital Design Implementation (NCSim)	0.9V	64.702
	1.1V	96.906

Table 3: Power Analysis of 4 bit MAC unit

Design	Supply Voltage	dcOp(static) (µW)	trans (µW)	transOp (µW)
4-bit MAC	0.9V	41.08	357.2	40.38
	1.1V	46.36	730.4	47.43

The above power analysis is done for the simulation rum time of 10ns.The static power is calculated by removing all sources except for the supply voltage. The dynamic power is calculated providing all the signals.

4. CONCLUSION

The 4 bit pipelined and clock gated MAC is implemented using Cadence Virtuoso in 90nm technology. Similarly the 2 bit pipelined and clock gated MAC architecture is implemented in both Cadence Virtuoso and Cadence NCSim (90nm). The two design approaches are compared to show the tremendous power reduction achieved in the handcrafted analog design. The PDP of the overall MAC unit is improved which can be used in all high speed and low power DSP applications. Also as the operation of the MAC is synchronized with the rising edges of the clock it is feasible to be used in real time applications of digital signal processors. The present work is also being extended to 8 bit unsigned and floating point architectures.

REFERENCES

[1] S. J. Jou, C. Y. Chen, E. C. Yang and C. C. Su, A pipeline Multiplier-Accumulator using a high speed, low power static and dynamic full adder design, IEEE custom Integrated circuit conference, 1995, pp. 593-596.

[2] West and Harris, CMOS VLSI Design: a circuits and systems perspective, Addison-Wesley Publishing Company, 3rd edition.

[3] N. Jithendra Babu, Rajkumar Sarma, A Novel Low Power and High Speed Multiply-Accumulate (MAC) Unit Design for Floating-Point Numbers, International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2015, pp. 411-417.

[4] C.S. Wallace, A Suggestion for a Fast Multiplier, IEEE Transactions on Electronic Computers, EC-13, 1964, pp. 14-17.

[5] Aditya Kumar Singh, Bishnu Prasad De, Santanu Maity, Design and Comparison of Multipliers Using Different Logic Styles, International Journal of Soft Computing and Engineering (IJSCE), Volume-2, Issue-2, May 2012.

[6] Avisek Sen, Partha Mitra, Debarshi Datta, Low Power MAC Unit for DSP Processor, International Journal of Recent Technology and Engineering (IJRTE), Volume-1, Issue-6, January 2013.

[7] Helga Evangelene, A Novel Low Power Hybrid Flipflop using Sleepy Stack Inverter Pair, Science and Information Conference 2014 August 27-29, 2014 London, UK.

[8] Jun Cheol Park and Vincent J. Mooney III, Sleepy Stack Leakage Reduction, IEEE transactions on very large scale integration (VLSI) systems, vol. 14, no. 11, November 2006.

[9] Partha Bhattacharyya, Bijoy Kundu, Sovan Ghosh, Vinay Kumar, Anup Dandapat, Performance Analysis of a Low-Power High-Speed Hybrid 1-bit Full Adder Circuit, IEEE Transactions On Very Large Scale Integration (VLSI) Systems, 2014.

[10] Naveen Kumar, Manu Bansal, Navnish Kumar, VLSI Architecture of Pipelined Booth Wallace MAC Unit, International Journal of Computer Applications (0975 – 8887) Volume 57– No.11, November 2012.