Journal of Scientific & Industrial Research
Vol. 79, February 2020, pp. 135–138

Low Latency Prefix Accumulation Driven Compound MAC Unit for Efficient FIR Filter Implementation

G Reddy Hemantha[1]*, S Varadarajan[2] and M N Giriprasad[1]

[1]Department of ECE, Jawaharlal Nehru Technological University Ananthapur, Ananthapuram

[2]Department of ECE, SVU College of Engineering, Sri Venkateswara University, Tirupati

Received 4 August 2018; revised 15 June 2019; accepted 8 October 2019

This article presents hierarchical single compound adder-based MAC with assertion based error correction for speculation variations in the prefix addition for FIR filter design. The VLSI implementation of approximation in prefix adder results show a significant delay and complexity reductions, all this at the cost of latency measures when speculation fails during carry propagation, which is the main reason preventing the use of speculation in parallel-prefix adders in DSP applications. The speculative adder which is based on Han Carlson parallel prefix adder structure accomplishes better reduction in latency. Introducing a structured and efficient shift-add technique and explore latency reduction by incorporating approximation in addition. The improvements made in terms of reduction in latency and merits in performance by the proposed MAC unit are showed through the synthesis done by FPGA hardware. Results show that proposed method outpaces both formerly projected MAC designs using multiplication methods for attaining high speed.

Keywords: Multiply Accumulate (MAC) unit, Distributed arithmetic, FIR filter, Compound adder

Introduction

High speed MAC units are inevitable in many real time applications like filter design, image processing, data acquisition and control[1-2]. Typical DSP applications need to carry out a large number of MAC operations as their implementation is based on multiplier kernels. It is investigated in many existing research works as arithmetic optimization models[2-4] have concludes that the design requires both high speed operations with significant improvements in complexity reduction. As mentioned[3], MAC computations in DSP applications need data path analysis in a large set. Targeting an optimized delay and area operators, techniques have been proposed[4]. The binary adder is used as the most fundamental building block in MAC unit, which performs not only additions but also divisions and multiplications. In order to diminish the delay and complexity to perform arithmetic operations of the adder several works have been done using various circumstances[7].

Speculative techniques

The speculative techniques[5-9] are extensively implemented as counter methods, which can also be used to diminish sub logarithmic delays by minimizing critical path of the required active input operands. Parallel prefix architectures are used[6] to perform accumulations with high speed in several digital devices. Recently, a several methods have developed[7].

High speed design

On the side the number of taps used in FIR filter design is keeps on increasing since it has been used in wide range of applications. In[8] Finite impulse response (FIR) design is used in Neural networks to extended and the process each neuron where synapse weights are replaced by finite impulse response filter for improved accuracy. Accumulator based radix-4 multiplier16 is used in RNS reverse converter system to carried out high speed FIR filter. This paper presents speculative prefix structure for Ling carry computation to attain both high speed and simple addition. As the critical path reduction methods had used for many pipelined architectures with no compromise in latency problems, with negligible hardware complexity. This work presents speculation by linking equations through parallel prefix computation with the core objective of variable Latency for the first time. The proposed multiplier unit is compared with the high radix Booth multiplier and high-speed

*Author for Correspondence
E-mail: hemanthag75@gmail.com

Vedic multiplier. In addition, the Distributed arithmetic method is used for FIR filter design to achieve high performance, in which one tap units are handled for all the bits within the bounded delay. Further, we are going to present speculation driven prefix topology based finite impulse response (FIR) filters design using residue number system (RNS) arithmetic. The aim of our work is to maximize the prefix accumulation in the application of RNS to the design high performance FIR filter design. To obtain this, we propose RAM based reverse conversion model followed by accumulation to produce the modular multiplication.

FIR Filter Design

Multiplier less technology yields numerous attractive features such as simplicity, regularity and modularity of architecture. In addition, with that, the DA technique can also produce high-speed implementation.

$$y[n]=\sum_{k=0}^{N-1} b_k \cdot x[n-k] \quad \ldots (1)$$

The output sequence of an N-tap finite impulse response (FIR) digital filter is computed as discrete time convolution as shown in Figure 1.

DA model

In general, multiplication and accumulation results separation are made by pipeline architecture with global clock routing. Our proposed FIR MAC contains only single adder, the partial product information is generated using zero detection unit followed by synchronized shift accumulation between PP and tap out from other level. The DA is computed with two parameters such as input signal component with unit delay elements (z) and FIR coefficient values (h) between neighbour delay buffers. Intel Quartus Prime is programmable logic device design software produced by Intel. Quartus prime enables analysis and synthesis of HDL designs, which enables the developer to compile their designs, perform timing analysis, examine RTL diagrams, simulate a design's reaction to different stimuli, and configure the target device with the programmer. Quartus Prime includes an implementation of VHDL and Verilog for hardware description, visual editing of logic circuits, and vector waveform simulation.

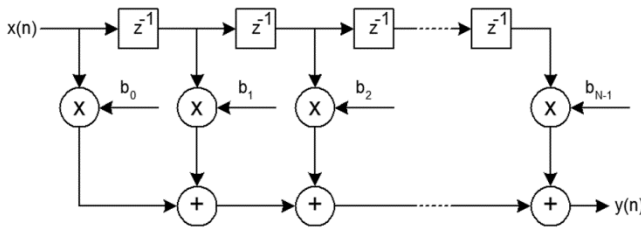

Fig. 1 — FIR MAC architecture

Parallel Prefix Computation

The overall cost required for reconfigurable architecture is optimized by exploiting the symmetrical measure of the main power switches with all intermediate power off modes. In pre and post processing units, Han Carlson prefix topology with the carry generation using single Brent Kung are used to achieve high speed parallel architectures with reduced critical path are employed.

Speculative prefix processing

In comparison with the conventional prefix architectures, the speculative computation is given the major variances in both speed and latency of accumulation. The approximated carry values are asserted and propagated to MSBs sub regions instead of waiting for the carry propagation from LSBs using the speculative computation. If the approximation fails with actual carries obtained from LSB side, the correction of the error is taken place by asserting the error correction is asserted causing any significant rise in critical path delay as shown in Figure 2.

Error Correction

This unit asserted when approximation fails to match with the actual carry signals which are propagated to MSBs regions called mis prediction. The actual addition output is attained by this unit, which is built with hierarchical levels of the EX-OR gates. The error correction unit of the proposed speculative parallel architecture is showed in the Figure 2. For any parallel prefix methods, the error correction can be done in the similar technique. It is

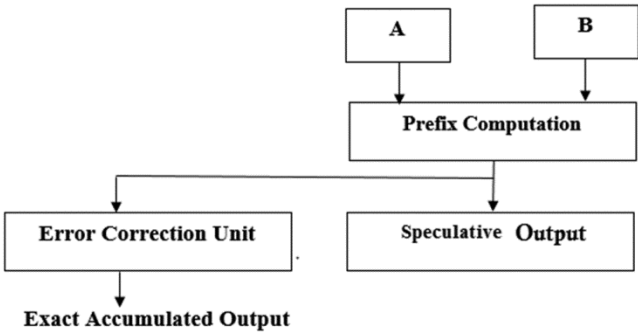

Fig. 2 — Proposed speculative parallel architecture

inevitable to get higher delay due to the integration of the error correction unit with basic cell structure of the speculative units. It shall cause an extra clock cycle for calculation in comparison with the speculative results with fine approximated carry values.

FIR Denoising Application

Here low pass FIR filter coefficient generation is carried out using FDA filter design analyzer tool and pre processed successfully using MATLAB. In addition to this, ECG signal is also generated and converted into 2's complement representation and these digital values are stored as a text file using block RAM. Memory unit basically consists of two units: 1) pre-processed FIR coefficients; 2) Noisy ECG signal component.

Pre-processing stage

In order to generate low pass filter coefficients filter design analysis tool is used with appropriate window selection methods and fine precision floating point values are exported in ASCII formatted text file using filter design analyzes tool. Similarly input noisy ECG signals are sampled and converted into numerical values. The signal samples and filter coefficients are converted into 16 bit binary values from MSB regions at the time of digital conversion. Finally, the approximated carries are stored and given to the sub block regions of the FIR filters.

Experimental Results

Here we compare the performance of the proposed MAC unit using single compound adder over existing well-known benchmark schemes explored in table 1 with the schemes described[9-10]. We extend this analysis using FIR filter implementation schemes and implement them for DSP applications as ECG de-noising process to verify the functionality of the FIR core as shown in Figure 3. Finally the hardware FPGA synthesis is carried without using any degree of EDA driven optimization since the objective of this work is to prove the performance of the aforementioned designs, using architectural level modifications to analyze the highest achievable complexity reduction and frequency. Here FIR filter designs are mapped to ALTERA cyclone III FPGA EP3C16F484C6 using Quartus II tool. As shown in table 2 speculative prefix accumulation driven FIR design outperformed its state-of-the-art accumulator based radix-4 FIR design[11] both in terms of hardware complexity as well as operating speed. As compared to lookup table based distributed arithmetic model both latency and design complexity is

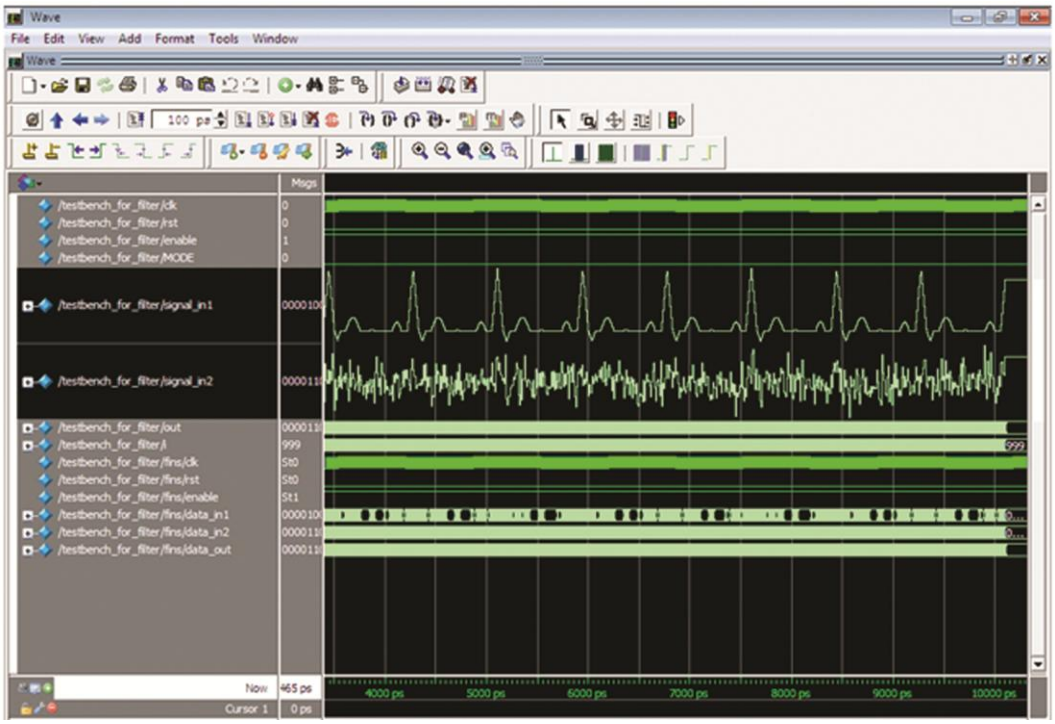

Fig. 3 — FIR filter simulation for ECG de noising

Table 1 — State-of-the-art comparison of MAC units using ALTERA CYCLONE II FPGA device.

Model used	LE's used	Fmax report(MHz)
Vedic model[9]	201	83
Booth+CSA model[10]	236	131.09
Proposed model	143	737.46

Table 2 — State-of-the-art comparison of 20 taps FIR filter units using ALTERA CYCLONE II FPGA EP2C35F672C6 device.

FIR model used	LE's used	Fmax report(MHz)	Power (mW)
Booth radix-4 CLA model [11]	5029	720.51	118.56
Proposed model	4893	730.12	115.23

greatly[12]optimized. Excluding the programmable shifters, our method consumes 28.85% and 39.40% lower than the existing methods as shown in Table 1. Our design also has a critical path delay of 1.35ns, which is several times faster than other MAC types. The efficiency could be consistent with nominal higher bit width operands. As it stands, the performance metrics of proposed bit serial shift-add based MAC network makes it a prerogative choice for DSP application.

Software

ALTERA cyclone III FPGA EP3C16F484C6 using Quartus II tool has been used. .Intel Quartus Prime is programmable logic device design software produced by Intel. Quartus prime enables analysis and synthesis of HDL designs, which enables the developer to compile their designs, perform timing analysis, examine RTL diagrams, simulate a design's reaction to different stimuli, and configure the target device with the programmer. Quartus Prime includes an implementation of VHDL and Verilog for hardware description, visual editing of logic circuits, and vector waveform simulation.

Conclusion

An analysis of FIR filter computation using single compound speculative adder is made in this paper to optimize arithmetic complexity and latency problems associated with filter design. A novel carry assertion is presented to reduce the partial speculative results in parallel prefix computation of inputs outputs. Using the proposed speculative formulation, a reconfigurable MAC architecture is derived for FIR filter design. A major aspect of the proposed architecture is that, it causes least latency to compute filter outputs with reduced complexity overhead. Moreover, the proposed method yields better delay reduction than the existing MAC units. FPGA synthesis result shows that the proposed method for single MAC size involves 5 times high speed and offer 30-40% hardware complexity reduction with the other methods.

References

1 Nikolaidis S, Karaolis E & Kyriakis-Bitzaros E D, Estimation of signal transition activity in FIR filters implemented by a MAC architecture, IEEE Trans Comput-Aided Design Integr Circuits Syst, **19** (1) (2000)164-169.

2 Amaricai A,Vladutiu M & Boncalo O, Design issues and implementations for floating-point divide–add fused, IEEE Trans Circuits Syst II, Exp Briefs, **57** (4) (2010) 295-299.

3 Peymandoust A & De Micheli G, Using symbolic algebra in algorithmic level DSP synthesis, Proc 38th annual Design Automation Conference, (2002) 277-282.

4 Jaberipur G & Gorgin S, An improved maximally redundant signed digit adder, *Comp Elect Engg*, **36** (3) (2010) 491-502.

5 Lu S L, Speeding up processing with approximation circuits, Computer, **37** (3) (2004) 67-73.

6 Koren I, Computer arithmetic algorithms, A K Peters, Natic, MA, 2nd Edition, (2002).

7 Brent R P, & Kung H T, A regular layout for parallel adders, *IEEE trans Comp*, **31** (4) (1982) 260-264.

8 Miljanović M Ninkov T Sušić, Z & Tucikesic S, Forecasting geodetic measurements using finite impulse response artificial neural networks, *Ind J Geo Marine Sci*, **46** (9) (2017), 1743-1750.

9 Tiwari H D Gankhuya G Kim C M & Cho Y B, Multiplier design based on ancient Indian Vedic Mathematics, So C Design IEEE Conference, **2**(2008) 11-65.

10 Daud NG N, Hashim F R, Mustapha M & Badruddin M S, Hybrid modified booth encoded algorithm-carry save adder fast multiplier, *Inform Comm Tech*, (2014)1-6.

11 Pari J B & Rani SP, Reconfigurable architecture of RNS based high speed FIR filter, *Ind J Eng & Mater Sci*, **21** (2) (2014), 233-240.

12 Jyothi G N & Sriadibhatla S, ASIC implementation of low power, area efficient adaptive fir filter using pipelined DA, *Electromagn and Telecommun*, (2019), 385-394.