International Journal of Hybrid Information Technology
Vol. 12, No. 2 (2019), pp.21-46
http://dx.doi.org/10.21742/ijhit.2019.7.1.04

Comparative Review of Floating-Point Multiplier

Marcus Lloyde George[1] and Geetam Singh Tomar[2]

[1]Dept. of Electrical and Computer Engineering, University of West Indies, St. Augustine, Trinidad and Tobago, India
[2]THDC Institute of Hydropower Engineering and Technology, Bhagirathipuram, Tehri 249 124, India
[1]marcus.george@sta.uwi.edu, [2]gstomar@ieee.org

Abstract

This paper presents a comprehensive comparative review of existing floating-point multiplier systems. The study focuses on single, double, quadruple and multi-precision floating-point multiplier architectures and seeks to identify engineering techniques involved in their development. A comparison of the performance of these systems in terms of metrics such as path delay, hardware utilization and even power consumption in some cases are carried out. Weaknesses in the systems reviewed along with possible gaps in the area of research are identified. This paper also serves to identify several recommendations and considerations for the development of a multi-precision floating-point multiplier system capable of treating the weaknesses of multiplier systems identified.

Keywords: *Arithmetic logic unit, Floating-point multiplier, Multi-precision floating-point, Multiplier system, Multiplier architecture, FPGAs in arithmetic*

1. Introduction

Arithmetic Logic Units (ALUs) are important components of processors that perform various arithmetic operations such as multiplication, division, addition, subtraction, cubing, squaring, etc. The ALU of some processors is divided into two units, an Arithmetic Unit (AU) and a Logic Unit (LU). Many times in Personal Computers (PCs) floating-point operations are performed by a floating-point unit that resides on a separate chip referred to as a numeric coprocessor. According to multiplication is the most elementary and most frequently used operation in ALUs. It allows one number to be scaled by another number.

Floating-point multiplication is the arithmetic operation most frequently utilized and is a very important component of many engineering applications such as signal processing, video processing, image processing, etc. The floating-point format can represent very small and large numbers when compared to fixed-point numbers, therefore the dynamic range of numbers that can be represented. Many operations and processes in science utilize floating-point arithmetic and therefore there is a need to develop units with shorter path delay, smaller hardware utilization and less power consumption [1][2][4].

[4] indicated that floating-point multiplication is the most commonly utilized operation in many applications involving digital signal processing. [4] further indicates that floating-point multiplication accounts for approximately 37% of the floating-point operations in benchmark

Article history:
Received (April 26, 2019), Review Result (June 11, 2019), Accepted (July 29, 2019)

Print ISSN: 1738-9968, eISSN: 2652-2233 IJHIT

applications. Multiplication consumes significant time compared to other arithmetic units used in basic mathematical computations [5]. Power management has also become extremely important especially in the case of portable electronic systems. Large power dissipation results in the chip having a higher temperature profile and as such, this affects the performance of the chip. According to the multiplier is a major power dissipation source and at the same time, high speed multiplication is a major requirement for high performance computing. As a result, it is beneficial in the area of mathematical computation to present faster and more efficient mechanisms for implementing mathematical operations which also can utilize less power. Before this can be achieved it may be useful to perform a comprehensive review of existing floating-point multiplier systems in such a way that can advise further evolution of the area of floating-point multiplication [6].

2. IEEE 754-2008 standard for floating-point numbers

Floating-point numbers are numbers that cannot be represented by integers because they contain fractional components or simply because they fall outside the range of values possible within a system's bit-width. By representing the numbers in floating-point format the accuracy and resolution of the numbers are preserved when compared to fixed-point format. In the binary system, floating-point numbers are represented in both single precision and double precision formats. Three components make up these formats: sign, exponent and mantissa components. The single precision floating-point format comprises a 1-bit sign (bit 31), 8-bit exponent (bits 23-30) and 23-bit mantissa (bits 0-22). The bias in this format is 127. The double precision floating-point format comprises a 1-bit sign (bit 63), 11-bit exponent (bits 52-62) and 52-bit mantissa (bits 0-51). The bias in this format is 1023. The Quadruple precision floating-point format comprises a 1-bit sign (bit 127), 15-bit exponent (bits 112-126) and 112-bit mantissa (bits 0-111). The bias in this format is 16383. The Octuple precision floating-point format comprises a 1-bit sign (bit 255), 19-bit exponent (bits 236-254) and 236-bit mantissa (bits 0-235). The bias in this format is 262,143. A summary of this data is given in [Table 1] [1][7][8][10]

Table 1. Comparison of single, double, quadruple and octuple precision floating-point formats

	IEEE-754 Standard Format			
	Single Precision	Double Precision	Quadruple Precision	Octuple Precision
Bits	32	64	128	256
Sign bit	1 (bit 31)	1 (bit 63)	1 (bit 127)	1 (bit 255)
Biased Exponent	8 (bits 23-30)	11 (bits 52-62)	15 (bits 112-126)	19 (bits 236-254)
Mantissa	23 (bits 0-22)	52 (bits 0-51)	112 (bits 0-111)	236 (bits 0-235)
Bias	127	1023	16383	262143
Range	2^{-126} – 2^{+127}	2^{-1022} – 2^{+1023}	2^{-16382} – 2^{+16383}	$2^{-262,142}$ – $2^{+262,143}$

Floating-point multiplication utilizing single, double, quadruple and octuple precisions formats first requires the computation of the sign bit which is done by the XOR operation of the two sign bits. The product exponent is computed by the summation of the two exponents. The value obtained is then added to the bias. The mantissa is computed by applying binary multiplication of the mantissa components of both floating-point numbers

3. Review of existing floating-point multiplier systems

Floating-point multiplication utilizing single, double, quadruple and octuple precisions formats first requires the computation of the sign bit which is done by an XOR operation of the sign bits of the input floating-point numbers. The product exponent is computed by summation of the two exponents, after which a bias is added. The value obtained is then added to the bias. The mantissa is computed by applying binary multiplication of the mantissa components of both input floating-point numbers. This section presents a review of existing implementations of floating-point multiplier systems [1].

3.1. Single precision floating-point multiplication

Single precision floating-point multiplication presented the development of a fast and compact GaAs IEEE single precision floating-point multiplier. In [9], Booth's algorithm is used together with a modified carry-save array to reduce the partial product addition and interconnection requirements indicated that GaAs is inherently superior in radiation hardness, saturation velocity, high temperature operation and electron mobility. The multiplier system consisted of exponent and mantissa block. The exponent block was responsible for the sign computation as well as the biased exponent computation. The system of [9] computed biased exponent simply by the addition of the biased exponents of each floating-point number after which the bias was subtracted, hence resulting in the biased exponent. The mantissa block of [9] is a 24x24-bit parallel multiplier consisting of an array of modified carry-save adders along with a final adder. The array of modified carry-save adders is used in the partial product reduction process to reduce the number of partial products from 24 to 13. A Wallace tree is then used to further reduce the partial products to 2. The final adder completes the partial product reduction process.

The multiplier system was designed using Vitesse H-GaAs-II 0.8um technology. Spice was used for the simulation of the performance of the system. When implemented on AT&T GaAsHFET 1um the delay of the system of [9]was 9.25ns, clock frequency was 74.75MHz, the area was 8x9.5mm^2 and power utilization was 7W. When implemented on GaAsMESFET 0.8um the delay of the system of [9] was 4ns, clock frequency was 14MHz, the area was 2.43x3.77mm^2 and power utilization was 3.5W.

[11] presented a single precision IEEE 754 floating-point multiplier with low power and high speed. [11] claimed that the bottleneck of any single precision floating-point multiplier was the 24x24 bit integer multiplier used in the mantissa calculation. In this implementation, the Urdhava Triyakbhyam algorithm that is a component of ancient Indian Vedic Mathematics was utilized for increasing the efficiency of implementation. Reconfiguration was also utilized at runtime for power savings. The system of [11] computed biased exponent simply by the addition of the exponents of each floating-point number after which the bias is added, hence resulting in the biased exponent. The mantissa multiplication component of [11] comprised of a 24x24-bit multiplication operation that was done using four parallel 12x12-bit multiplication modules. These 12x12-bit multiplication units were constructed using a series of 4x4-bit multiplier units. As such the entire 24x24-bit multiplication operation was divided into thirty-

six (36) 4x4-bit multiplier units operating in parallel. The system of [11] was implemented on the Xilinx Virtex E XCV300e package BG432, speed grade -8 using Xilinx Webpack 6.1. Two versions of the system were implemented – with and without reconfiguration. [11] claimed that the proposed multiplier without reconfiguration had an area utilization of 2967 slice registers and an estimated delay of 37.553ns, while the proposed multiplier reconfiguration had an area utilization of 3149 slice registers and an estimated delay of 41.203ns.

[12] presented hardware description of IEEE 754 single precision floating-point multiplier. Analysis was made using Booths algorithm and Canonical Signed Digit (CSD) algorithm was used in the implementation. The system of [12]computes the biased exponent by the sum of the two exponents of the input floating-point numbers followed by the addition of the bias to the sum of both exponents. The mantissa multiplication was done by use of the Booth Algorithm and CSD. The system was synthesized for the Virtex E XCV300e, package bg432, speed grade -8, using Xilinx ISE after which it was simulated using Modelsim. [12] indicates that the proposed multiplier achieved a delay of 22.859 ns, an area of 1331 slice registers and a maximum frequency of 333.33MHz. [12] claimed that the system implemented resulted in an improvement of 57.77% in the area utilized and 44.52% in the system delay when compared to the system implemented in [11].

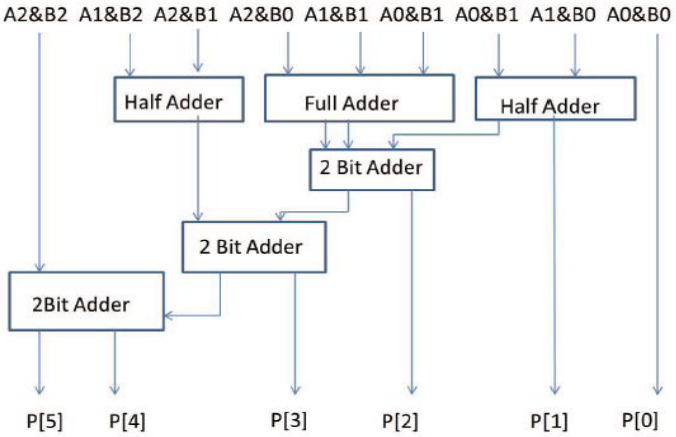
[13] proposed a new reversible design for the single precision floating-point multiplier by use of an operand decomposition technique [13]. The system performed the 24x24 bit reversible multiplication using nine (9) reversible 8x8-bit Wallace tree multipliers. A new reversible design of the 8x8 bit Wallace tree multiplier which was optimized in terms of quantum cost, delay and number of garbage outputs was proposed by [13] The system of [13] computed the biased exponent by first subtracting bias from the exponent of the multiplicand, then adding this result to the exponent of the multiplier input. [13] computed performed mantissa multiplication by use of a 24x24 reversible partition multiplier. There was no mention of the delay or hardware utilization of the system.

[14] presented an efficient implementation of the IEEE 754 single precision floating-point multiplier unit. The system of [14] computed biased exponent simply by the addition of the biased exponents of each floating-point number using a binary adder after which the bias was subtracted, hence resulting in the biased exponent. The mantissa multiplication operation of the system of [14] was done using a 24x24 carry-save multiplier which consisted of three stages that were constructed using carry-save adders. The system was implemented in VHDL using Xilinx ISE and Precision Synthesis tools MG (2010), the target being Xilinx Virtex-5 5VFX200TFF1738 using a timing constraint of 300MHz. [14] indicates the rounding was not implemented in the system, however, a normalizer was included in the system. The system was simulated to obtain area and maximum frequency. Its parameters were compared with that of the single-precision implementation found in the Xilinx Coregen. [14]did not indicate the delay or hardware utilization of the system. Based on the comparison it was realized that the system implemented in [14] utilized grea]ter area than that of the system from Xilinx Coregen, while the system of [14] has a greater max frequency (301.114MHz) than that of the unit from Xilinx Coregen (221.484MHz).

[16] presented an efficient floating-point multiplier system using the Karatsuba algorithm and the IEEE 754 format for floating-point numbers. The system was implemented using Verilog HDL on the FPGA cyclone II device as the target and Altera-Quartus II as the development environment. [16] incorporated a three-stage pipelining scheme with latency 8 clock cycles in the design. The system of [16] computed biased exponent simply by the addition of the biased exponents of each floating-point number using ripple carry adders after which the bias was subtracted, hence resulting in the biased exponent. The mantissa multiplication

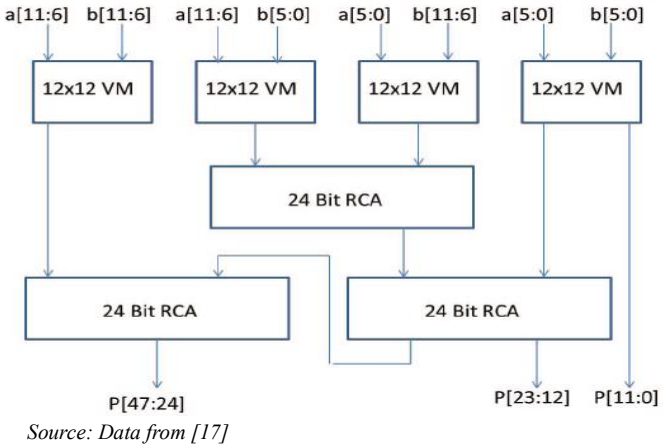
operation of[16] was done using a 24-bit unsigned multiplier, Karatsuba Multiplier which utilizes Vedic multiplication. The multiplier system of [16] was simulated using Modelsim. The max frequency of operation was determined to be 77.434MHz while the delay was determined 12.92 ns.

[17] presented the development of a 24-bit Vedic multiplier using a 3x3 Vedic multiplier as the basic building block. [17] also proposed implementation of an IEEE-754 single precision floating-point multiplier unit capable of handling rounding, overflow and underflow conditions. The proposed and conventional floating-point units are implemented and simulated using the ISE simulation tool. The system implementation was done on iWave using the Spartan 6 XC6S1x25t-2fgg484 as the development platform. The system of [17] computed biased exponent simply by the addition of the biased exponents of each floating-point number using a binary adder after which the bias is subtracted, hence resulting in the biased exponent. The system implemented by[17] utilizes 3x3 Vedic multiplier blocks (seen in Figure I). A 6x6 multiplier block is constructed using 3x3 blocks after which a 12x12 block is constructed using 6x6 blocks. Finally, a 24x24 multiplier block (seen in [Figure 2]) is constructed using 12x12 Vedic multiplier blocks and three 24-bit ripple carry adders [18].



Source: Data from [17]

Figure 1. 3x3 vedic multiplier block implementation



Source: Data from [17]

Figure 2. 24x24 Vedic multiplier block implementation

The system developed in [17] and conventional multipliers were compared in terms of maximum combinational path delay and hardware utilization. Results indicate that the method utilized by [17] had a speed that was 21.7% faster than that of the conventional system (seen in [Table 2]). The area on the Spartan 6 utilized for the system of [17] is also smaller than that used by the conventional system [Table 3]. The system also utilized 1018 sliced LUTs and 96 bonded IOBs as seen in [Table 3].

Table 2. Path delay comparison for proposed and conventional systems

Proposed Floating Point Multiplier (ns)	Conventional Floating Point Multiplier (ns)	Increased Speed (%)
49.797	63.595	21.7

Source: Data from [17]

Table 3. Area comparison for proposed and conventional systems

Parameter	Proposed Floating Point Multiplier	Conventional Floating Point Multiplier
Number of Slice LUTs	1018	1347
Number of Bonded IOBs	96	96

Source: Data from[17]

[7] presented the design and implementation of a floating-point multiplier to reduce path delay [Figure 3]. Reduction of the path delay can be done by reducing the delay caused by the propagation of the carry in adders utilized in the multiplier design [7]. The architecture implemented was based on the IEEE-754 standard for single precision format. The system of [7] computed biased exponent simply by the addition of the biased exponents of each floating-point number using a binary adder after which the bias is subtracted, hence resulting in the biased exponent. The system of [7] carried out mantissa multiplication simply by first computing partial products using radix-4 booth programming, after which a Wallace tree was used for partial product reduction.

The modules were implemented in Verilog and synthesized for the Spartan 3 XC3S500-4GFG320 device using Synopsys Design Compiler, and a pipelined approach was utilized. The system was later interfaced with a DSP processor. The system of [7] computed biased exponent simply by the addition of the biased exponents of each floating-point number after which the bias was subtracted, hence resulting in the biased exponent. The multiplier unit implemented in [7] and the performance of the system along with all submodules (Modified Booth Encoder, KoggeStone Adder, Pipelined KoggeStone Adder and Wallace Tree) were summarized. These four systems were compared in terms of power, area, path delay and hardware utilization and the results are shown in [Tables 4] and [Table 5].

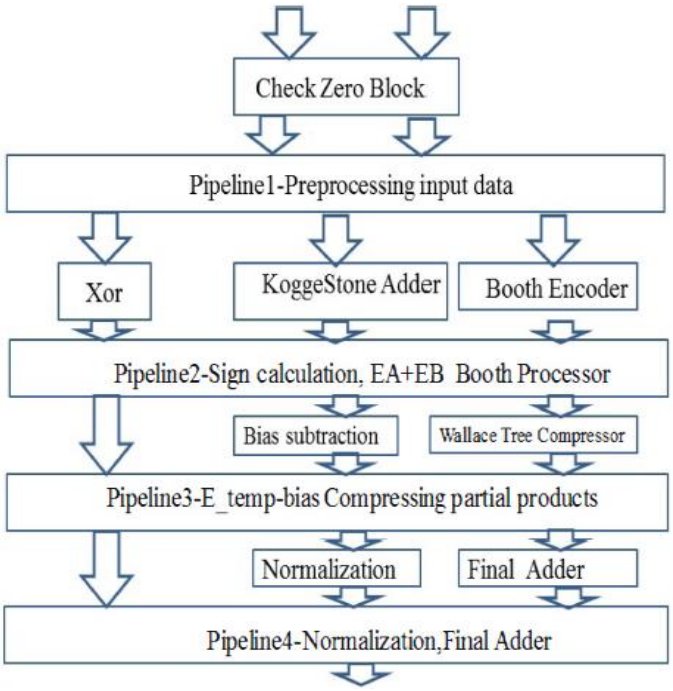


Figure 3. Proposed multiplier unit

Table 4. Performance summary of multiplier sub-units

	Power (µW)	Area (µm^2)	Delay (ns)	No. of LUTs	No. of Slices
Proposed Multiplier Unit	173.72	13325.45	29.72	2270 /9312	1208 /4656
Modified Booth Encoder	21.72	355.35	4.84	125 /9312	72 /4656
KoggeStone Adder	27.05	360.00	3.72	355 /9312	196 /4656
Pipelined KoggeStone Adder	37.25	420.00	3.92	358 /9312	208 /4656
Wallace Tree	44.25	524.20	7.92	895 /9312	515 /4656

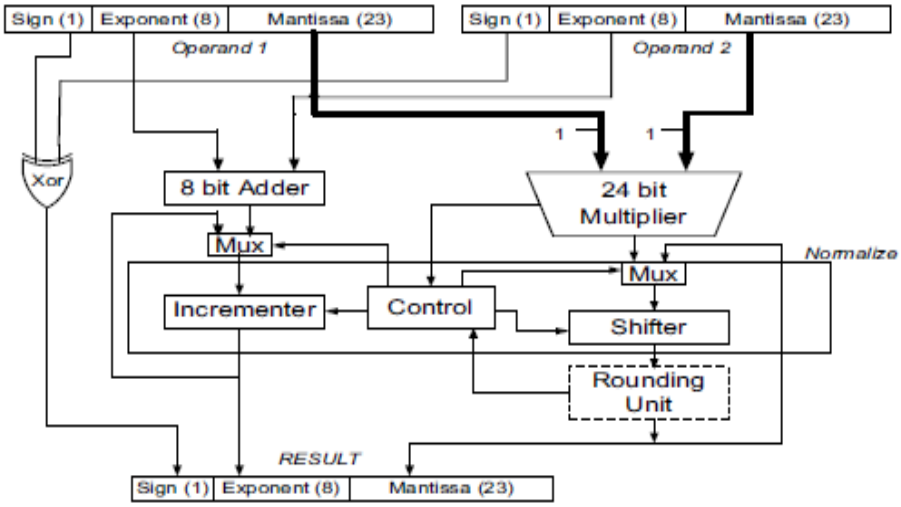
Source: Data from [7]

Table 5. Performance summary of the proposed multiplier

	Power (µW)	Area (µm^2)	Delay (ns)	No. of LUTs	No. of Slices
Proposed Multiplier Unit	173.72	13325.45	29.72	2270 /9312	1208 /4656

Source: Data from [7]

[19] presented the design and implementation of a low power probabilistic floating-point multiplier [Figure 4]. In [19] probabilistic computation is utilized to attain large energy savings in the floating-point multiplier at the expense of calculation errors/accuracy.



Source: Data from [19]

Figure 4. Architecture of 32-bit single precision floating-point multiplier

[19] indicated that the 24-bit multiplier block of a single precision floating-point multiplier consumes 81% of the power consumed by the floating-point multiplier block. [19] also indicated that 18% of the power consumed is done by the rounding component. In [19] rounding is not utilized; hence this means that 99% of the power is consumed by the 24-bit multiplier block. According to [19] AND gates account for 5% of the power consumed by the 24-bit multiplier block, while full adders account for the remaining 95%. Since most of the power is utilized by full-adders, [19] applied a low power approach to the design and implementation of the full-adders.

[19] presented two methods that could have been applied to making the 24-bit multiplier operate at low power. The first termed as the Sleep Technique allows some of the less significant full adder logic of the multiplier to go to sleeping mode while the remaining stay awake while operating at the normal voltage level. The second method utilizes BIased Voltage Scaling (BIVOS) where the full-adders of a 24-bit multiplier are provided with a biased supply voltage depending on the significance of the computation being carried out. [19] proposed a new technique for low power operation which utilized both Sleep and BIVOS techniques. When starting from the columns of least significance, some columns are switched to sleeping mode while the remaining are supplied with a biased voltage. The columns of lesser significance with which are not currently in sleep mode are operated at low voltage when compared to the columns of greater significance [19]. Results of [19] indicate that implementing the floating-point multiplier using the BIVOS+Sleep strategy results in a power saving of 62% as indicated in [Table 6].

Table 6. Comparison of image quality vs power consumption for bivos, sleep and bivos+sleep floating-point multiplier

PSNR (dB) Relative to the Ideal Image	Energy Relative to the Base Case		
	BIVOS	Sleep	BIVOS + SLEEP
58 dB	80%	75%	66%
47 dB	55%	51%	38%

Source: Data from [19]

[20] presented a self-timed 32-bit floating-point multiplier with a carry-lookahead adder. The implementation was based on the IEEE 754 32-bit floating-point multiplier standard. [20] also presented a self-timed carry-lookahead adder that was implemented using dual-rail signaling [21] in the input, sum and carry bits. The sign bit was generated from an XOR operation of sign bits of both inputs. The biased exponent was obtained by summation of both biased exponents of inputs using the self-timed carry-lookahead adder followed by subtraction of the bias. The mantissa multiplication operation was done by addition and shifting operations. The addition operations were done using the self-timed carry-lookahead adders constructed in [20]. The system was implemented using Xilinx ISE 14.4 and performance was compared with that of a Synchronous implementation. [20] claimed that the new system saw marginal improvements in path delay compared to synchronous implementation. [20] also claimed that the use of the self-timed floating-point multiplier developed resulted in a 20% improvement in power consumption when compared to the synchronous multiplier implementation.

3.2. Double precision floating-point multiplication

[23] presented the implementation of high speed floating-point double-precision multiplier which is compliant with the IEEE 754 standard for floating-point numbers, hence handling overflow, underflow, and rounding conditions. The system of [23] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. [23] broke up the input mantissa bits into ten (10) parts, partial products were generated, after which partial products were summed together. The system of [23] was implemented on the Virtex-6 xc6vlx75t-3ff484 using Xilinx ISE 12.1i. The system was simulated using Modelsim 6.6c. The minimum period was determined to be 2.411ns which maximum frequency was determined to be 414.714MHz. The results of the implementation of [23] were compared with that of single precision implementations from [14] and Xilinx Core from Xilinx Coregen. Area utilized and a maximum frequency of [23] was determined to be greater than that of [14] and Xilinx Core from Xilinx Coregen.

[24] indicated that the double-precision floating-point multiplier required a 52x52 mantissa multiplication operation and as such have proposed a novel approach towards the decrease of this huge mantissa multiplication operation. [24] utilized the Karatsuba and Urdhva Tiryagbhyam techniques in the implementation of the multiplier system. [24] indicated that traditionally the partial products of the multiplier are added separately and took a significant amount of time to complete. The proposed method presented by [24] concurrently added the partial products during the multiplication operation, hence reducing the delay. The double-precision floating multiplier system of [24] was implemented in Verilog HDL on Xilinx ISE using Virtex-5 FPGA as the target. The system of [24] catered for the detection of all the exceptional cases of the IEEE standard such as overflow, underflow and infinite zero.

Normalization is also incorporated in the implementation. The system of [24] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. [24] claims that the path delay of the Karatsuba multiplier was determined to be 18.139 ns, while the path delay of the Urdhva Tiryagbhyam multiplier was determined to be 15.034 ns. Hardware utilization was determined to be (798 sliced registers, 1378 sliced LUTs) and (525 sliced registers, 936 sliced LUTs) respectively [24]

[25] presented an area efficient runtime reconfigurable double precision floating-point multiplier architecture. The system conformed to the IEEE-754 floating-point standard. [25]presented three multiplier architectures for double precision floating-point multiplication. The first architecture (Truncated Block Multiplication - TBM) is a truncated multiplication block. This system was designed using Vedic mathematics. In this system, the two 53-bit mantissae are multiplied together after which the result is trimmed back to 53-bits by suitable truncation or rounding of the result. This was expected to result in a minor loss of accuracy [25]. The second architecture (3-Partition Karatsuba Multiplication - PKM) was included for regaining the loss of accuracy from the multiplication operation performed by the first block. In this architecture, both 53-bit mantissae are separated into three parts – two 18-bit and one 17-bit component. Each component of both numbers is multiplied and the partial products are generated and added. The three results are then added and the overall result is brought back to standard form by post normalization and rounding methods [25]. The third architecture (Double Precision Dual Single Precision Multiplier - DPdSP) was responsible for performing either single or double precision floating-point operation. The design was achieved via a resource-sharing approach. The system can produce one double precision floating-point result of two single precision floating-point results [25]. The system of [25] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. All three systems developed in [25] were implemented in VHDL using Xilinx ISE and simulated using the Synopsys tool. The target devices for this research were the Virtex-4 xc4vfx100-12ff1517 and the Virtex-5 xc5v1x155-3ff1760. Performance Comparison for the three architectures is presented in [Table 7].

Table7. Performance comparison for TBM, PKM and DPDSP

Performance Parameters	Target Device	TBM	PKM	DPdSP
Number of Slice Registers	Virtex 4	1923	3419	3288
	Virtex 5	2439	4468	4484
Number of 4-Input LUTs	Virtex 4	3348	5950	5729
	Virtex 5	-	-	-
Delay / ns	Virtex 4	44.91	46.74	58.93
	Virtex 5	40.03	42.23	49.91

Source: Data from[25]

3.3. Quadruple precision floating-point multiplication

[30] indicated that quadruple precision multiplication is required especially in high precision requirements of a given application. The system of [30] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of

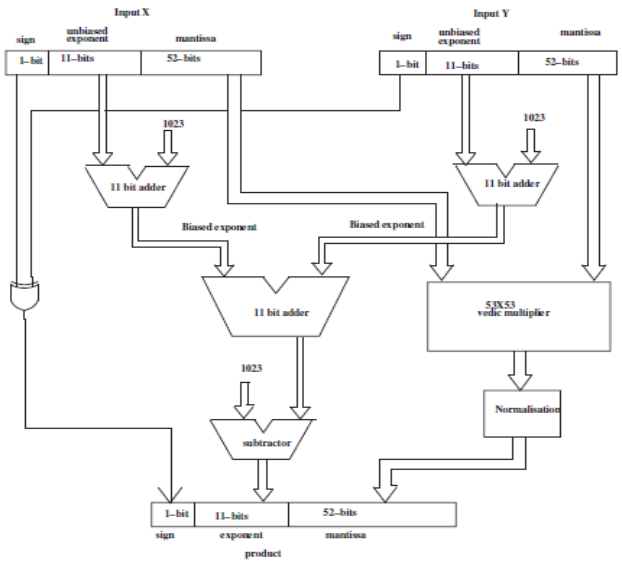
the bias. In the case of the mantissa multiplication procedure, [30] proposed a hardware efficient approach for the implementation of fully pipelined integer multipliers. [30] later presented the implementation of quadruple precision floating-point multiplier which utilized DSP48 as the basic building blocks. Block sizes of 17-bits were used as the starting point. These were used in constructing bigger multiplier blocks such as 34-bit, 51-bit, 66-bit, 130-bit and 113-bit multiplier blocks. The 113-bit multiplier block was utilized in mantissa multiplication for this system. The systems were implemented in Verilog HDL using Xilinx ISE. The target utilized was the Xilinx Virtex 4 xc4vfx100-12ff1517. The implemented system was simulated using Modelsim-SE. [30] indicated that the proposed 113-bit binary multiplier and quadruple floating-point multiplier had hardware utilization of 2373 and 2464 cycles respectively. [30] also indicated that the proposed 113-bit binary multiplier and quadruple floating-point multiplier had max frequencies of 310MHz each. It was unclear whether the latency given for both was in nanoseconds or cycles.

[26] presented a new quadruple precision floating-point multiplication algorithm that was based on polynomial multiplication. The system of [26] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. The mantissa multiplier using the polynomial multiplication method required the system to utilize 98 multipliers and 1598 LEs if a multiplier core from the development environment is used. The new system was then implemented in Verilog HDL using Quartus II 8.1 IDE. The target for this implementation was the Altera EP2C7089C6. Results of [26] indicated that the multiplication algorithm used effectively reduced hardware resource occupancy. [26] also indicated that the delay of the floating-point multiplier, in this case, is 96 ns.

3.4. Multi-precision floating-point multiplication

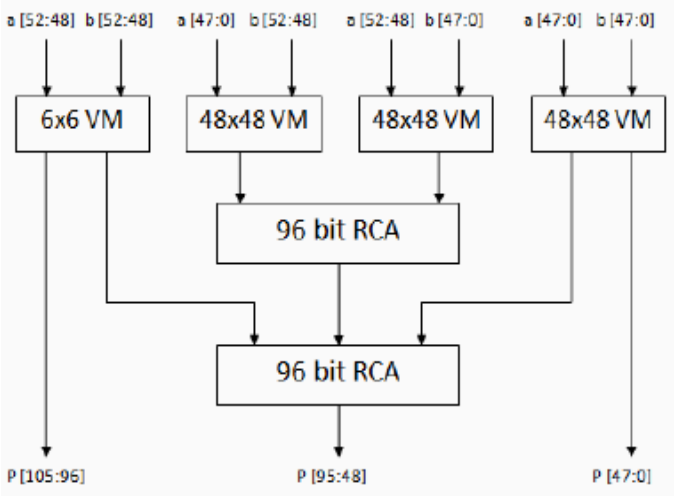
[1] presented a Vedic multiplier based on the IEEE 754 which was capable of both single and double precision floating-point multiplication with performance comparable to the Karatsuba and Booth type floating-point multipliers [Figure 5]. The system of [1] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. The mantissa multiplication component of [1] was developed by utilization of Vedic multiplier principles and as such all bits of both inputs were subjected to vertical and crosswise multiplication. The floating-point multiplier of [1] was synthesized for the Virtex-7 7v2000tflg1925-2. According to [1] the application of the Vedic multiplier to digital signal processing reduced path delay by 40-60% when compared to conventional procedures. The findings of [1] conclude that the Vedic multiplier utilized less hardware and also has a shorter path delay than both the Karatsuba floating-point multiplier. The Vedic multiplier utilized more hardware than the Booth type floating-point multiplier and had a shorter delay. [1] also concluded that Booth type is a poor choice for single-cycle multiplication but can be efficiently implemented in the case of highly pipelined systems.

[27] attempted to improve the performance of the multiplier developed by [17], hence producing a multiplier that was capable of facilitating both IEEE 754 single and double precision floating-point multiplication. Implementation was done using Vedic Mathematics using the same approach outlined in [17]. The system of [27] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. The mantissa multiplication component of the single precision multiplier was implemented using a 24x24 Vedic Multiplier while the double precision mantissa multiplication was done using a 53x53 Vedic Multiplier as shown in [Figure 6].



Source: Data from [1]

Figure 5. Block diagram of IEEE-754 double precision floating-point multiplier



Source: Data from [27]

Figure 6. Proposed 53x53 Vedic multiplier block implementation

The system was implemented using Xilinx ISE 14.6. [Table 8] presents the area utilization for the double precision floating-point multiplier of [27] while [Table 9] presents the delay comparison for the previous system from [17] and proposed system from [27].

Table 8. Area utilization for double precision floating-point multiplier

Logic utilization	Used	Available	Utilization
No. of slices LUTs	5153	204000	2%
No. of bonded IOBs	192	600	32%

Source: Data from [27]

Table 9. Delay comparison for previous system and proposed systems

Delay	Used	Available
SP FPM	49.497 ns	21.823 ns
DP FPM	-	45.169 ns

Source: Data from [27]

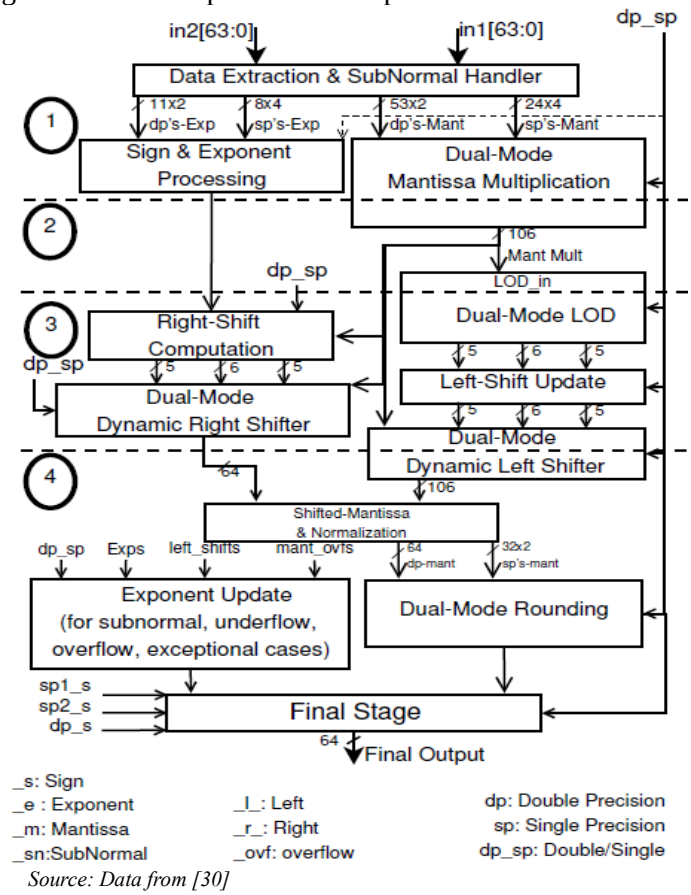
[4] presented a pipelined IEEE floating-point multiplier system that was capable of carrying out either single-precision or double precision multiplication. [4] claims that the latency of the system in single precision operation was 2 cycles while in double-precision operation it was 3 cycles. The clock cycles however are not related to the reference clock but rather to the clock common to all pipelined registers. Hardware requirements for this system were minimized by the use of a half-sized multiplication array and by also ensuring that both precisions use the same rounding units. [4] indicated that the system presented supported all IEEE rounding modes. [4] presented a new rounding algorithm that supplied different precisions, hence allowing its use by multiple precisions.

[28]presented a hardwired algorithm for the computation of variable precision floating-point multiplication based on the use of a parallel multiplier of size *m* which would be used to compute n×m bits. [28] only focused on the mantissa component of the floating-point inputs. A very important component of the implementation of [28] is that the partial products are added as soon as they are computed to reduce demands on memory for storing partial products. [28] the variable precision is brought about by the use of both floating-point and fixed-point formats. Resources and the resulting architecture are dedicated to the multiplication of operands of size ranging from 1x64 to 64x64 bits with the period of $n^2 \times 33$ns. The system of [28] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. The system of [28] was implemented on Virtex 2 XC2V1000 (-5) bg575 using Xilinx ISE 6.3. The system was simulated using Modelsim PE 6.0. Results of [28] indicated that the path delay for the system was 19,008 ns in single precision mode and 135,168 ns in double precision mode. The system also occupied 2381 slice registers on the Virtex 2 platform.

[29] presented a multi-mode floating-point multiplier system which operated with the single, double and quadruple precision formats specified by IEEE 754-2008 standard. The system of [29] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. The system of [29] is pipelined to maximize throughput, hence allowing the execution of one quadruple precision floating-point multiplication, two double precision floating-point multiplications processes and four single precision floating-point multiplication processes. The system was implemented in VHDL and synthesized for a 45NM technology using Synopsys for the front-end implementation and Cadence for the back-end. [29] indicated that the proposed system was implemented using a divide and conquer technique which utilized high precision multiplications by performing smaller sized multiplications and at the end adding up the partial products to produce the final result. [29] performed an example implementation of the system on VLSI and as a subsequent verified that the implementation achieved a maximum operating frequency of 505MHz.

[30] presented a configurable dual-mode floating-point multiplier which was capable of functioning as double precision mode as well as the process in single precision mode (see Figure VII). The architecture proposed is based on the flow of floating-point multiplication capable of processing in both normal and sub-normal operands along with exceptional case

handling. The system proposed by [30] was implemented for ASIC (UMC 90nm) technology. [30] utilized state of the art computational flow of floating-point multiplication in the design of the proposed system. Each stage was reconstructed using "efficient resource sharing and tuned datapath" [30] as a means of minimizing the system's multiplexing circuitry. The system is also pipelined to maximize throughput and is aimed at ASIC UWMC 90nm technology. The system is comprised of four pipeline stages. The first pipeline stage dealt with data extraction, sign and exponent processing and dual-mode mantissa multiplication (using a Dadda-tree multiplier and Kogge-Stone adder). The first 6 levels of the Dadda-tree multiplier lay in the first pipeline stage while the last 2 levels fell in the second pipeline stage. The system of [30] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. After implementation, the system was tested. The system implemented in [30] utilized 11.6% more area and 6.74% more time in its execution when compared to a single mode double-precision multiplier.



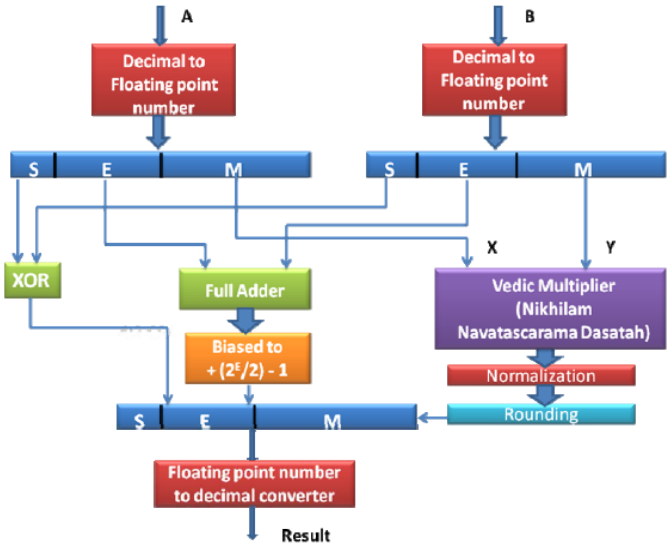
Source: Data from [30]

Figure 7. DPdSP multiplier architecture

[2] presented a run-time reconfigurable floating-point multiplier system which was implemented on the Xilinx Virtex 4 using Xilinx ISE 14.7. [2] utilized a combination of the Urdhva-Tiryagbhyam algorithm and Karatsuba algorithm to implement the unsigned binary multiplier, to further increase the efficiency of the multiplier. This multiplier utilizes crosswise and horizontal multiplication operations. The multiplier system implemented in [2]had 6 modes of operation which are selected based on the requirements of the application to which it must be used. Therefore, the system can perform floating-point multiplication for varying mantissa

sizes all depending on the precision requirements of the application. The system of [2] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. Results of [2] indicate that the proposed 32-bit binary multiplier had a delay of 13.141 ns while the proposed single-precision floating-point multiplier had a delay of 16.392 ns.

[31] presented the design and implementation of a multi-precision floating-point multiplier using Vedic mathematics as indicated in [Figure 8]. The resulting multiplier supports single, double and quadruple precision floating-point multiplication. The system of [31] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. The mantissa multiplication component is based on Vedic mathematics which utilizes crosswise and vertical multiplication operations. The results of [31] indicated that the multi-precision multiplier presented utilized slightly more area than the double-precision floating-point multiplier.



Source: Data from [31]

Figure 8. Block diagram of the resulting universal multiplier system

[32] presented a multi-function double precision floating-point multiplier system capable of performing one double-precision multiplication or one vector multiplication of two 2D vectors in single-precision floating-point format. The system of [32] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. The system of [32] utilized multiplexers for the selection of data in the two modes – double precision and single-precisions mode. The binary multiplication section utilizes 4:2 Compressor Tree logic for partial product reduction. Pipelined and non-pipelined versions are produced and compared with a conventional floating-point multiplier in terms of latency, area and power consumption. Results of tests conducted by [32] as seen in [Figure 9] indicate that the latency of the proposed system is 16% less than that of the conventional floating-point multiplier and the latency of the pipelined version of the proposed system is 24% less than that of the conventional floating-point multiplier. [32] also reported that the area of the proposed system is 47% smaller than that of the conventional floating-point multiplier and the area of the pipelined version of the proposed system is 49% smaller than that of the conventional floating-point multiplier. [32] finally reported that the proposed system consumes 6% more

power than the conventional floating-point multiplier and the pipelined version of the proposed system consumes 16% more power than the conventional floating-point multiplier.

LATENCY COMPARISON

Latency		32	64	Dual-64	[7]0.25μm	[8]90nm
COMB	ns	1.10	1.45	1.60	---	2.59
	FO4	34	45	49	---	58
PIPE	ns	0.70	0.90	0.98	5.07	---
	FO4	21.5	27.7	30	40.6	---

AREA COMPARISON

Area		32	64	Dual-64	[7]0.25μm	[8]90nm
COMB	μm^2	13587	52097	56569	---	253500
	gates	9435	36178	39284	---	176042
PIPE	μm^2	11745	41336	45571	---	---
	gates	8156	28076	31646	35103	---

POWER OF THE PROPOSED DESIGN

Power(mW)	32	64	Dual-64	[7]0.25μm	[8]90nm
COMB	6.71	24.00	25.41	--	212
PIPE	8.05	25.68	29.78	--	--

Source: Data from Liu et al. (2015, 60)

Figure 9. Comparison of latency, area and power consumption for proposed systems and conventional system

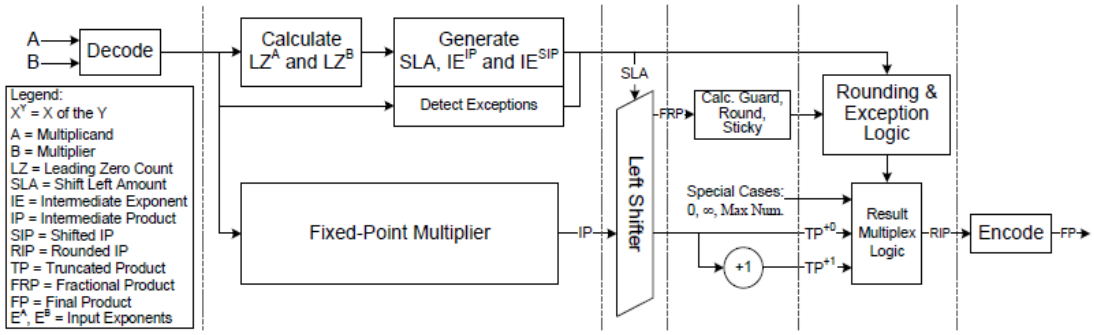
[33] presented an FPGA-based iterative hardware architecture for quadruple precision floating-point multiplication which is also capable of processing single, double and double-extended precision data. [33] utilized a series expansion method of division along with wide integer multiplication to further optimize the FPGA implementation. Some expansion equations of [33] were implemented: equation (1) of [33] was implemented using adders, subtractors and multipliers. For the multi-precision component, a unified expression (3) of [33] was utilized for supporting all four precision computations. The system in this case utilized a look-up table of size 2^8 x 113 along with a multiplier of size 114x114-bit which is used iteratively with an FSM control unit. The system of [33] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias. The system of [33] was demonstrated on Xilinx Virtex-7 FPGA and according to [33]the implementation yielded improvement in latency along with significant saving in the area utilized when compared to [34] as shown in [Table 10].

[35] presented a fully parallel decimal floating-point multiplier based on parallel fixed-point multiplier [36] and which was compliant with the 2007 draft of the IEEE P754 standard for floating-point arithmetic (see [Figure 10]). [35] claimed that the novelty in the design was that at the time, it was the first parallel decimal floating-point multiplier which offered low latency and high throughput. The multiplier system in [35] used a 64-bit decimal floating-point format with significands encoded in Densely Packed Decimal (DPD) [37]. The system had 16 mantissa bits and a bias of 398.

Table 10. Performance comparison between multipliers produced

	(Diniz and Govindu 2006)	(Jaiswal and So 2016)
Latency	118 (QP)	9/11/12/13 (SP/DP/DPE/QP)
Throughput	NA (QP)	8/10/11/12 (SP/DP/DPE/QP)
LUTs	26811	7440
FFs	13809	2584
DSP48	-	18
Freq (MHz)	50	89

Source: Data from[33]



Source: Data from [35]

Figure 10. Block diagram of the decimal floating-point multiplier system

The radix-10 parallel fixed-point multiplier utilized special BCD digit recoding for the reduction of the logic utilized by the system[35]. The system allowed for the generation of partial products in parallel [35]. The fixed-point multiplier design consisted of three components – generation of multiplicand multiples, partial product selection, and partial product reduction [35]. The multiplier of [35] converted the floating-point number to BCD format and then utilized the fixed-point multiplier to generate a 32-bit BCD number referred to as an intermediate product (IP). The fixed-point multiplier utilized by [35] generated sixteen (16) decimal partial products in parallel and then computed the sum of these partial products using a carry-save adder (CSA) tree. The 32-bit intermediate product was then shifted to form the shifted immediate product (SIP). The SIP was then utilized in the computation of the final result of the decimal floating-point multiplier system [35]. The system of [35] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias.

The system of [35] was implemented in Verilog on LSI Logic's gflxp 0.11um CMOS standard cell library and Synopsys Design Compiler Y-2006.06-SPI. The performance summary of the parallel decimal floating-point multiplier is given in [Table 11].

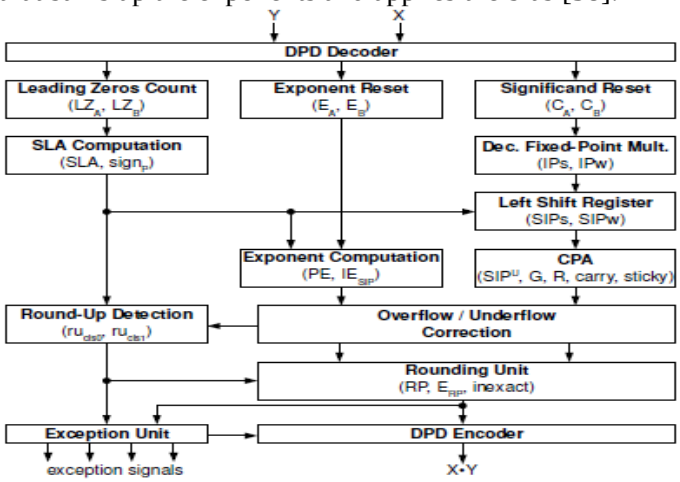
Table11. Performance comparison of parallel decimal floating-point multiplier

Component	Delay Breakdown		Area Breakdown	
	ps	%	µm2	%
Fixed-point Multiplier	2990	67.3%	326,841	50.2%
Round Logic	810	18.2%	14,125	2.2%
Left Shift	280	6.3%	17,414	2.7%
Datapath	360	8.1%	26,605	4.1%
Interconnect	N/A	N/A	270,243	41.5%
Entire Design	4400	%	651,435	100%

Source: Data from [35]

[38] presented an IEEE 754-2008 compliant parallel decimal floating-point multiplier system capable of exploiting the features of the Virtex-5 FPGA (see [Figure 11]). The system of [38] implemented early estimation of the shift-left amount, SLA and efficient decimal rounding technique. The system of [38] also provided all required exception handling, rounding modes, overflow and gradual underflow. It also incorporated pipelining to increase the throughput of the system. The system is fully combinational and based on fast partial product generation, BCD recording schemes and BCD-4221 carry-save adder (CSA) reduction tree [38].

The system first begins by decoding the input operands to extract the floating-point components – sign, biased exponent and mantissa [38]. The system then performs decimal fixed-point multiplication on the mantissa to give a result. The product sign is produced by XOR operation on the input signs. The intermediate exponent (IE) is calculated by an exponent computation unit that sums up the exponents and applies the bias [38].

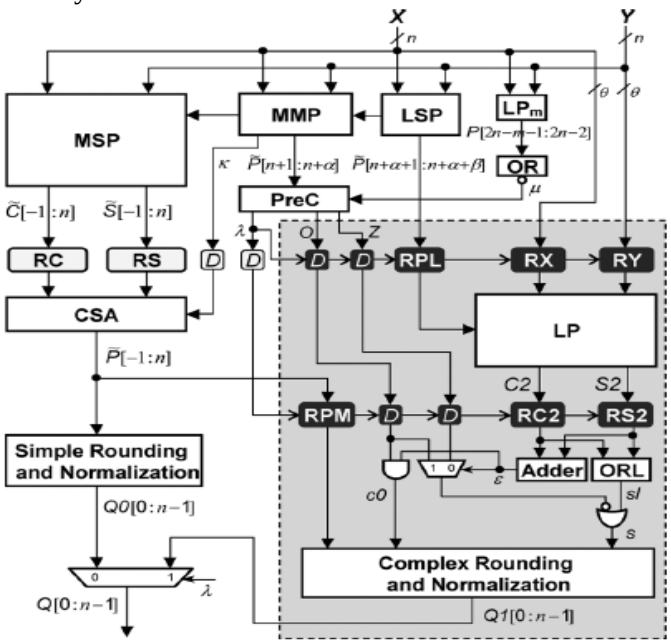


Source: Data from [38]

Figure 11. Block diagram of the parallel decimal floating-point multiplier

The system was implemented on Xilinx ISE for the target Xilinx Virtex5 FPGA device XC5VLX110T. [38] claimed that the developed systems achieved decimal floating-point multiplication within 35ns and an operating frequency of 192 MHz using 13 pipeline stages. [38] also indicated that the hardware utilization was approximately 8000-9000 LUTs.

[38] proposed a variable-latency floating-point multiplier architecture that was compliant with IEEE 754-1985 and was deemed suitable for low-power applications (see [Figure 12]). The multiplier architecture of [39] splits the mantissa multiplier into the upper and lower components, and predicts the sticky bit, carry bit, and mantissa product from the upper part. If the prediction is correct, the computation of the lower part is disabled and the rounding operation is simplified, hence allowing the system to consume less power [38] The system of [39] computed the biased exponent by the summation of both biased exponents of inputs using binary adders followed by subtraction of the bias.



Source: Data from [39]

Figure 12. Block diagram of proposed low-power floating-point multiplier

The proposed multiplier was implemented in Verilog and synthesized using Synopsys Design Compiler with TSMC CMOS standard cell library. Synopsys NanoSim and Cadence SOC Encounter and were used to perform placement, routing, and full transistor-level power simulation [39]. Power simulation was performed at a clock frequency of 33.3 MHz and 1.8 V with 10,000 random input patterns. [39] indicated that results show that the proposed multiplier saved power and energy at the expense of area and delay overheads [39]. The performance summary for the low-power floating-point multiplier is given in [Table 12].

Table12. Performance summary for low-power floating-point multiplier

n	Multiplier	Area (μm2)	Delay (ns)
53	Fast_FM VL_FM (α, m)	548138 574268	7.15 8.00
24	Fast_FM VL_FM (α, m)	150416 153039	4.52 5.20

Source: Data from [39]

3.5. Summary of performance of various floating-point multipliers

The reviewed multipliers of sections II(a) - II(d) were compiled and provided in [Tables 13-15]. Some sources reviewed did not provide all the required information and as such those sections in the tables were left unfilled. It is important to note also that no existing implementations of Octuple-Precision Floating-point multiplication were found.

Table 13. Summary of various single precision floating-point multipliers reviewed

Source	Path Delay / ns	Hardware Utilization (Slice Registers)	Maximum Frequency / MHz
[9] – Compact GaAs	4.00		
[11] – without reconfiguration – Xilinx Virtex E (XCV300E-8BG432)	37.55	2,967	
[11] – with reconfiguration – Xilinx Virtex E (XCV300E-8BG432)	41.20	3,149	
[12] – Xilinx Virtex E (XCV300E-8BG432)	22.86	1,331	333.33
[28] – Xilinx Virtex 2 (XC2V1000-5BG575)	n2x33 per cycle	2,381	
[14] – Xilinx Virtex 5 (XC5VFX200TFF1738)			301.11
[16] – Altera Cyclone II	12.92		77.43
[17] – Xilinx Spartan 6 (XC6S1X2ST-2FGG484)	49.80	1,018	
[7] – Xilinx Spartan 3 (XC3S500 – 4GFG320)	29.72	1,208	
[1] – Xilinx Virtex 7 (7V2000TFLG1925-2)	28.02	2,928	
[2] – Xilinx Virtex 4 (XC4VFX12-10FF668)	16.39	977	226.51
[24] – Karatsuba Version – Xilinx Virtex 5 (XC5VLX110T-1FF1136)	18.14		
[24] – Urdhva Tiryagbhyam Version – Xilinx Virtex 5 (XC5VLX110T-1FF1136)	15.03		
[27] – Xilinx Virtex 6 (nomenclature not specified)	21.82	2,153	
[33] – ASIC (UMC 90nm)	9.00	7,440	89.00

Table 14. Summary of various double precision floating-point multipliers reviewed

Source	Path Delay / ns	Hardware Utilization (Slice Registers)	Maximum Frequency / MHz
[2] – Xilinx Virtex 4 (XC4VFX12-10FF668)	18.97	3,877	173.95

[1] – Xilinx Virtex 7 (7V2000TFLG1925-2)	34.08	15,494	-
[25] – TBM – Xilinx Virtex 4 (XC4VFX100-12FF1517) – Xilinx Virtex 5 (XC5VLX155-3FF1760)	44.91	2,439	-
[25] – PKM – Xilinx Virtex 4 (XC4VFX100-12FF1517) – Xilinx Virtes 5 (XC5VLX155-3FF1760)	46.70 42.23	4,468	-
[25] – DPdSP – Xilinx Virtex 4 (XC4VFX100-12FF1517) – Xilinx Virtes 5 (XC5VLX155-3FF1760)	58.93	4,484	-
[27]– Xilinx Virtex 6 (nomenclature not specified)	45.17	2,153	-
[33] – ASIC (UMC 90nm)	11.00	7,440	89.00

Table 15. Summary of various quadruple precision floating-point multipliers reviewed

Source	Path Delay / ns	Hardware Utilization (Slice Registers)	Maximum Frequency / MHz
[34]– Xilinx Spartan 3 (nomenclature not specified)	118.00	26811	50.00
[22] – Xilinx Virtex 4 (XC4VFX100-12FF1517)		2464	310.00
[26]– Altera EP2C7089C6	96.00		
[33]– Xilinx Virtex 7 (nomenclature not specified)	13.00	7440	89.00

4. Considerations for development of novel floating-point multiplier system

The development of a novel floating-point multiplier system capable of eliminating (or at least minimizing) the effects of the weaknesses of existing multiplier systems is a viable consideration for the benefit of digital electronic systems.

All existing systems reviewed in this paper utilized biased exponent calculators which computed the product biased exponent by the summation of both biased exponents of both inputs using binary adders followed by subtraction of the bias using a subtractor or by twos' complement subtraction using adders. All existing systems catered for biased exponent calculation in the case of single precision only or double precision only or, quadruple precision only or, both single and double precision modes only. None has catered for the biased exponent calculation for all four precision modes: single, double, quadruple and octuple precision modes. None of them are capable of computing multiple batches of biased exponents for the various

precision modes. The development of a novel multi-precision biased exponent calculator capable of supporting single, double, quadruple and octuple-precision modes and also capable of computing the biased exponent of multiple batches of input floating-point numbers in single, double, quadruple-precision modes (hence increasing the overall system throughput) is a viable consideration.

Most of the multiplier systems reviewed in this paper carried out the processes of partial product generation, partial product storage and partial product reduction. For example, multipliers developed in [1],[40] and [41] perform partial product reduction using Wallace or Dadda multipliers, thereafter the results are compressed using compressors. Others like [3] [15]use a combination of multiplier and compressor techniques to perform the partial product reduction segment. Most of the existing systems reviewed utilized Vedic mathematics for partial product generation. Multiplier systems in [5] and [11] developed Vedic multipliers by utilizing smaller multipliers as building blocks to developing bigger multipliers.

Some systems like in [19] proposed a technique for low power operation which utilized both Sleep and BIVOS techniques. When starting from the columns of least significance, some columns are switched to sleeping mode while the remaining is supplied with a biased voltage. This method resulted in a loss in accuracy. Other systems such as the multiplier architecture of [39] split the mantissa multiplier into the upper and lower components, and predicted the sticky bit, carry bit, and mantissa product from the upper part. If the prediction was correct the computation of the lower part was disabled and the rounding operation was simplified, hence allowing the system to consume less power. There is a need for development of a multiplier system that is capable of accumulating partial products as they are generated, hence reducing the delay at the expense of hardware utilization. At the same time, there is no implementation of the binary multiplier that can be utilized for all four floating-point multiplier modes – single, double, quadruple and octuple precisions modes. Only a few such as the multiplier of [27] and others catered for multi-precision and at best processed 2 batches of input single precision floating-point numbers or 1 in double precision mode. The development of such a system is also a viable consideration. In [42][43] authors have demonstrated hardware design and modification in the binary system design, which has resulted in less simulation time.

Finally, none of the existing binary multiplication systems reviewed analyzed past multiplication operations to further reduce the path delay of the multiplication operation. Focusing on previous multiplication operations could benefit future multiplications, hence preventing the system from having to undergo lengthy partial product generation operations especially in the case of quadruple and octuple precision modes where the number of partial products can become very large. The inclusion of a novel system called Mantissa Similarity Investigation (MSI) in the floating-point multiplier system which can capable of further reduction in path delay is also a viable consideration.

5. Conclusions

This paper presented a comprehensive comparative review of existing floating-point multiplier systems. The study focused on single, double, quadruple and multi-precision floating-point multiplier architectures and identified engineering techniques involved in their development. A comparison of the performance of these systems in terms of metrics such as path delay, hardware utilization and even power consumption in some case were carried out. Weaknesses in the systems reviewed along with possible gaps in the area of research were identified. This paper also served to identify several recommendations and considerations for the development of a multi-precision floating-point multiplier system capable of treating the

weaknesses of multiplier systems identified. The development of systems involving such considerations is expected to result in shorter path delay than all existing implementations of floating-point multiplication reviewed in this paper. These contributions will likely be extremely useful to arithmetic operations in digital and computer systems presently and in the future.

References

[1] Kodali, Ravi Kishore, Lakshmi Boppana, and Sai Sourabh Yenamachintala, "FPGA implementation of vedic floating-point multiplier," IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), 19-21 February, pp.1-4, New York: IEEE, (**2015)** DOI: 10.1109/SPICES.2015.7091534

[2] Arish S. and R. K. Sharma, "Run-time reconfigurable multi-precision floating-point multiplier design for high speed, low-power applications," 2nd International Conference on Signal Processing and Integrated Networks (SPIN), 19-20 February, pp.902-907, New York: IEEE, **(2015)** DOI: 10.1109/SPIN.2015.7095315

[3] Arish S. and R. K. Sharma, "An efficient binary multiplier design for high speed applications using karatsuba algorithm and urdhva-tiryagbhyam algorithm," Global Conference on Communication Technologies (GCCT), 23-24 April, pp.192-196, New York: IEEE, **(2015)** DOI: 10.1109/GCCT.2015.7342650

[4] Even G., S. M. Mueller, and P. M. Seidel, "A dual mode ieee multiplier," Proceedings of 2nd Annual IEEE International Conference on Innovative Systems in Silicon, 8-10 October, pp.282-289, New York: IEEE, **(1997)** DOI: 10.1109/ICISS.1997.630271

[5] Sharma, Richa, Manjit Kaur, and Gurmohan Singh, "Design and FPGA implementation of optimized 32-Bit vedic multiplier and square architectures," International Conference on Industrial Instrumentation and Control (ICIC), 28-30 May, pp.960-964, New York: IEEE, **(2015)** DOI: 10.1109/IIC.2015.7150883

[6] Anitha P. and P. Ramanathan, "A new hybrid multiplieusing dadda and wallace method," International Conference on Electronics and Communication Systems (ICECS), 13-14 February, pp.1-4, New York: IEEE, **(2014)** DOI: 10.1109/ECS.2014.6892623

[7] Sunesh N.V and P Sathishkumar "Design and implementation of fast floating-point multiplier unit," International Conference on VLSI Systems, Architecture, Technology and Applications (VLSI-SATA), 8-10 January, pp.1-5, New York: IEEE, **(2015)** DOI: 10.1109/VLSI-SATA.2015.7050478

[8] IEEE (Institute of Electrical and Electronic Engineers), 754-2008-IEEE Standard for Floating-Point Arithmetic, Revision of ANSI/IEEE Std 754-1985, New York: IEEE, **(2008)**

[9] Cui Xiaoping, Weiqiang Liu, Xin Chen, Earl Swartzlander, and Fabrizio Lombardi, "A modified partial product generator for redundant binary multipliers," IEEE Transactions on Computers, vol.65, no.4, pp.1165-1171, **(2015)** DOI: 10.1109/TC.2015.2441711

[10] Huntsman C. and D. Cawthron, "The MC68881 floating-point coprocessor," IEEE Micro, vol.3, no.6, pp.44-54, **(1983)** DOI: 10.1109/MM.1983.291185

[11] Thapliyal, Himanshu, and M. B. Srinavas, "A novel time-area-power efficient single precision floating multiplier," Proceedings of MAPLD 16-18 June, pp.1-3, New York: IEEE, **(2005)**

[12] Siddamal, Saroja V., R. M. Banakar, and B. C. Jinaga, DELTA 2008. "Design of high-speed floating-point multiplier," 4th IEEE International Symposium on Electronic Design, Test and Applications, 23-25 January, pp.285-289, New York: IEEE, **(2008)** DOI: 10.1109/DELTA.2008.19

[13] Nachtigal, Michael, Himanshu Thapliyal, and Nagarajan Ranganathan, "Design of a reversible single precision floating-point multiplier based on operand decomposition," 10th IEEE Conference on Nanotechnology (IEEE-NANO), 17-20 August, pp.233-237, New York: IEEE, **(2010)** DOI: 10.1109/NANO.2010.5697746

[14] Al-Ashrafy, Mohamed, Ashraf Salem, and Wagdy Anis, "An efficient implementation of floating-point multiplier," Saudi International Electronics, Communications and Photonics Conference (SIECPC), 24-26 April, pp.1-5, New York: IEEE, **(2011)** DOI: 10.1109/SIECPC.2011.5876905

[15] MG (Mentor Graphics), Precision Synthesis User's Manual. Ottawa: MG, **(2010)**

[16] Mehta, Anand, C. B. Bidhul, Sajeevan Joseph, and P. Jayakrishnan, "Implementation of single precision floating-point multiplier using karatsuba algorithm," International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), 3-5 November, pp.254-256, New York: IEEE, **(2013)** DOI: 10.1109/ICGCE.2013.6823439

[17] Paldurai K. and K. Hariharan, "FPGA Implementation of Delay Optimized Single Precision Floating-Point Multiplier," International Conference on Advanced Computing and Communication Systems, 5-7 January, pp.1-5, New York: IEEE, **(2015)** DOI: 10.1109/ICACCS.2015.7324094

[18] Mano M. Morris and Charles R. Kime, Logic and Computer Design Fundamentals. New Jersey: Prentice Hall, **(1997)**

[19] Gupta, Aman, Satyam Mandavalli, Vincent J. Mooney, Keck-Voon Ling, Arindam Basu, Henry Johan, and Budianto Tandianus, "Low power probabilistic floating-point multiplier design," 2011 IEEE Computer Society Annual Symposium on VLSI, 4-6 July, pp.182-187, New York: IEEE, **(2011)** DOI: 10.1109/ISVLSI.2011.54

[20] Beohar, Salty, and Sandip Nemade, "VHDL implementation of self-timed 32-bit floating-point multiplier with carry look ahead adder," International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 25-27 May, pp.772-775, New York: IEEE, **(2016)** DOI: 10.1109/ICACCCT.2016.7831743

[21] Cheng, Fu-Chiung, Stephen H. Unger, Michael Theobald, and Wen-Chung Cho, "Delay-insensitive carry-look ahead adders," Proceedings of 10th International Proceedings VLSI Design, Conference, 4-7 January, pp.37-63, New York: IEEE, **(1997)**

[22] Jaiswal, Manish Kumar, C. Ray, and C. Cheung, "Area-efficient FPGA implementation of quadruple precision floating-point multiplier," IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW), 21-25 May, pp.376-382, New York: IEEE, **(2012)** DOI: 10.1109/IPDPSW.2012.46

[23] Ramesh, Addanki Purna, A. V. N. Tilak, and A. M. Prasad, "An FPGA-based high speed IEEE-754 double precision floating-point multiplier using verilog," International Conference on Emerging Trends in VLSI, Embedded System, Nano Electronics and Telecommunication System (ICEVENT), 7-9 January, pp.1-5, New York: IEEE, **(2013)** DOI: 10.1109/ICEVENT.2013.6496575

[24] Rao, Y. Srinivasa, M. Kamaraju, and D. V. S. Ramanjaneyulu, "An FPGA implementation of high speed and area efficient double-precision floating-point multiplier using urdhva tiryagbhyam technique," Conference on Power, Control, Communication and Computational Technologies for Sustainable Growth (PCCCTSG), 11-12 December, pp.271- 276, New York: IEEE, **(2015)** DOI: 10.1109/PCCCTSG.2015.7503923

[25] Shanmugapriyan S. and K. Sivanandam, "Area efficient run time reconfigurable architecture for double precision multiplier," IEEE 9th International Conference on Intelligent Systems and Control (ISCO), 9-10 January,pp.1-6, New York: IEEE, **(2015)** DOI: 10.1109/ISCO.2015.7282355

[26] Lei Kang and Yan Xiao-Ying, "Design and implementation for quadruple precision floating-point multiplier based on fpga with lower resource occupancy," 5th International Conference on Intelligent Systems Design and Engineering Applications (ISDEA), 15-16 June, pp.326-329, New York: IEEE, **(2014)** DOI: 10.1109/ISDEA.2014.80

[27] Havaldar, Soumya, and K S Gurumurthy, "Design of vedic IEEE 754 floating-point multiplier," IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). 20-21 May, pp.1131-1135, New York: IEEE, **(2016)** DOI: 10.1109/RTEICT.2016.7808008

[28] Anane N., H. Bessalah M. Issad, and M. Anane, "Hardware implementation of variable precision multiplication on FPGA," 4th International Conference Design & Technology of Integrated Systems in Nanoscal Era, 6-9 April, pp.77-81. New York: IEEE, **(2009)** DOI: 10.1109/DTIS.2009.4938028

[29] Manolopoulos K., D. Reisis, and V. A. Chouliaras, "An efficient multiple precision floating-point multiplier," 18th IEEE International Conference on Electronics, Circuits and Systems (ICECS), 11-14 December, pp.153-156, New York: IEEE, **(2011)** DOI: 10.1109/ICECS.2011.6122237

[30] Jaiswal, Manish Kumar, and Hayden K. H. So., "Dual-mode double precision / two-parallel single precision floating-point multiplier architecture," IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC), 5-7 October, pp.213-218, New York: IEEE, **(2015)** DOI: 10.1109/VLSI-SoC.2015.7314418

[31] Mangalath, Nithu. S., R. Arokia Priya, and P. Malathi, "An efficient universal multi-mode floating-point multiplier using vedic mathematics," International Conference on Advances in Communication and Computing Technologies (ICACACT), 10-11 August, pp.1-4, New York: IEEE, **(2014)** DOI: 10.1109/EIC.2015.7230724

[32] Liu De, Mingjiang Wang, Yiwen Wang, and Hang Su, "A multi-functional floating-point multiplier," IEEE 9th International Conference on Anti-counterfeiting, Security, and Identification (ASID), 25-27 September, pp.56-60, New York: IEEE, **(2015)** DOI: 10.1109/ICASID.2015.7405661

[33] Jaiswal, Manish Kumar, and Hayden K. H. So, "Architecture for quadruple precision floating-point division with multi-precision support," IEEE 27th International Conference on Application-specific Systems, Architectures and Processors (ASAP), 6-8 July, pp.239-240, New York: IEEE, **(2016)** DOI: 10.1109/ASAP.2016.7760807

[34] Diniz P. and G. Govindu, "Design of a field-programmable dual-precision floating-point arithmetic unit," International Conference on Field Programmable Logic and Applications, 28-30 August, pp.1-4, New York: IEEE, **(2006)** DOI: 10.1109/FPL.2006.311302

[35] Hickman, Brian, Andrew Krioukov, and Michael Schulte, "A Parallel IEEE P754 decimal floating-point multiplier," 25th International Conference on Computer Design, 7-10 October, pp.56-62, New York: IEEE, **(2007)** DOI: 10.1109/ICCD.2007.4601916

[36] Vazquez A., E. Antelo, and P. Montuschi, "A new family of high–performance parallel decimal multipliers," 18th IEEE Symposium on Computer Arithmetic (ARITH '07), 25-27 June, pp.195-204, New York: IEEE, **(2007)**

[37] Cowlishaw Mike, "Densely packed decimal encoding," IEE Proceedings – Computers and Digital Techniques vol.149, no.3, pp.102-104, **(2002)**

[38] Baesler, Malte, Sven-Ole Voigt, and Thomas Teufel, "An IEEE 754-2008 decimal parallel and pipelined fpga floating-point multiplier," International Conference on Field Programmable Logic and Applications. 31 August-2 September, pp.489-495, New York: IEEE, **(2010)** DOI 10.1109/FPL.2010.98

[39] Kuang, Shiann-Rong, Jiun-Ping Wang, and Hua-Yi Hong, "Variable-latency floating-point multipliers for low-power applications," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol.18, no.10, pp.1493-1497, **(2010)** DOI: 10.1109/TVLSI.2009.2025167

[40] Abraham, Sumod, Sukhmeet Kaur, and Shivani Singh, "Study of various high speed multipliers," International Conference on Computer Communication and Informatics (ICCCI), 8-10 January, pp.1-5, New York: IEEE, **(2015)** DOI: 10.1109/ICCCI.2015.7218139

[41] Vyas, Keerti, Ginni Jain, Vijendra K. Maurya, and Anu Mehra, "Analysis of an efficient partial product reduction technique," International Conference on Green Computing and Internet of Things (ICGCIoT), 8-10 October, pp.1-6, New York: IEEE, **(2015)** DOI: 10.1109/ICGCIoT.2015.7380417

[42] George, Marcus, and Geetam Singh Tomar, "Hardware design procedure: principles and practices," 5th International Conference on Communication Systems and Network Technologies, 4-6 April, pp.834-838, New York: IEEE, **(2015)** DOI: 10.1109/CSNT.2015.198

[43] GS Tomar and Marcus George, "Modified binary multiplier architecture to achieve reduced latency and hardware utilization," Springers Wireless Personal Communication, vol.98, no.4, pp.3554-3561, **(2018)**

Authors

Marcus Llyode George

Received the B. Sc. degree in Electrical and Computer Engineering from the University of the West Indies, St. Augustine in 2007, his MPhil degree in Electrical and Computer Engineering from the University of the West Indies, St. Augustine in 2011 and his PhD degree in Electrical and Computer Engineering from the University of the West Indies, St. Augustine in 2019 under the guidance of Prof. Stephen Gift and Prof GS Tomar. He is the author of several academic books. His research engineering interest include the business administration, strategic planning and management, engineering education, formal specification, modelling and verification, field programmable architectures, intelligent electronic instrumentation and biomedical engineering.

Prof. Geetam Singh Tomar

Received his UG, PG, and Ph. D. degrees in electronics engineering from The Institute of Engineers, Calcutta, REC Allahabad and State Technology Univ Bhopal, respectively and Post Doctoral Fellowship from University of Kent, Canterbury U.K. He is Director of THDC Institute of Hydropower Engg & Technology, Tehri (Govt of Uttarakhand), Prior to this worked as visiting Professor at School of Computing, University of Kent, UK, Faculty at University of west indies, Trinidad & Tobago. Apart from this, has served in IIITM Gwalior, MITS Gwalior and other Institutes of repute. Received International Pluto award for academic achievements in 2009 from IBC, Cambridge UK.

He is IEEE Senior Member, Fellow IETE and IE (I) member CSI and ISTE. He is actively involved in research and consultancy in the field of Air Interface and Advanced communication networks. He is chief editors of 5 International Journals and for 08 in past. Has published more than 190 research papers in international journals/conferences and has written 09 books from Springer, CRC, IGI and indian publishers and 04 patents. He is actively involved in IEEE activities and has organized 30 IEEE International conferences in India and other countries with which he is associated.

Hi is member, Working group setup to finalise IEEE Standard; ISO/IEC/IEEE 21451-1-4. Member of Technical committee IEEE SMC and Industrial Electronics. Associated with many boards and committees of Government and currently working for prototyping of the research outcomes in Sensor networks and cognitive radio networks.