High Performance, Low Latency Double Digit Decimal Multiplier on ASIC and FPGA

Rekha K. James, K. Poulose Jacob

Department of Computer Science
Cochin University of Science and Technology
Kochi, Kerala, INDIA
rekhajames@cusat.ac.in, kpj@cusat.ac.in

Sreela Sasi

Department of Computer and Information Science
Gannon University
Erie, PA, USA
sasi001@gannon.edu

Abstract— **Decimal multiplication is an integral part of financial, commercial, and internet-based computations. This paper presents a novel double digit decimal multiplication (DDDM) technique that offers low latency and high throughput. This design performs two digit multiplications simultaneously in one clock cycle. Double digit fixed point decimal multipliers for 7digit, 16 digit and 34 digit are simulated using Leonardo Spectrum from Mentor Graphics Corporation using ASIC Library. The paper also presents area and delay comparisons for these fixed point multipliers on Xilinx, Altera, Actel and Quick logic FPGAs. This multiplier design can be extended to support decimal floating point multiplication for IEEE 754-2008 standard.**

Keywords- Decimal Multipliers, High Performance, ASIC, FPGA, Carry Save Adders

I. INTRODUCTION

The majority of the world's commercial and financial data is stored and manipulated in decimal form. A simple example is a pocket calculator, which is based on some form of decimal arithmetic. Currently, general purpose computers do decimal computations using binary arithmetic. Binary data can be stored efficiently and manipulated very quickly on two-state computers. Additional points favoring binary arithmetic include better error characteristics and less hardware to implement the same function. However, there are compelling reasons to consider decimal arithmetic, particularly for business computations. The reasons include human's natural affinity for decimal arithmetic and the inexact mapping between some decimal and binary values. Binary floating-point values can only approximate certain common decimal numbers. For example a value of 0.1 requires an infinitely recurring binary pattern of zeros and ones. When an average user performs a calculation such as addition of 0.1 and 0.9, the expected result is 1.0. The user would find it very confusing to be presented with an answer of 0.999999. In this world of precision, such errors generated by conversion between decimal and binary formats are no more tolerable. In many cases, the law requires that results generated from financial calculations performed on a computer exactly match those carried out using pencil and paper. That can be achieved only if the calculations are executed in decimal. Recently, support for decimal arithmetic has received increased attention due to this growing importance in financial analysis, banking, tax calculation, currency conversion, insurance, telephone billing and accounting which cannot tolerate such errors.

Due to the increasing significance of decimal arithmetic, standard specifications are recently added to the draft revision of the IEEE 754 Standard for Floating-Point Arithmetic [7]. The new IEEE 754-2008 standard defines a single data type that can be used for integer, fixed-point and floating-point decimal arithmetic. Hardware support for decimal operations, however, has been limited. But the scenario is set to change with the cost of die space continually dropping and the significant speedup achievable in hardware [3]. But till now, there is little in the way of hardware assist for financial applications that perform operations on data stored in decimal form. This is because decimal arithmetic operations are typically more complex, slower and occupy more area leading to more power and less speed when implemented in hardware. Hence, the major consideration while implementing decimal arithmetic is to enhance its speed and reduce area as much as possible. General-purpose processors, such as those from AMD and Intel, provide the ability to add and subtract values stored in decimal format. More-complex operations like multiplication and division must be constructed from the ground up using shifts, addition and subtraction. To speed up such engineering and scientific calculations, today's computers include high-performance, floating-point coprocessors. Nowadays, Field Programmable Gate Arrays (FPGAs) are frequently used for complex designs that are oriented for functioning as co-processors.

Decimal multipliers are typically implemented using an iterative approach because of their complexity.

Several iterative designs for fixed-point decimal multiplication have been proposed [1, 6, 10, 11] in the seventies and eighties. These designs iterate over the digits of the multiplier, and based on the value of the current digit, either successively add the current multiplicand or a multiple of the multiplicand. The multiples are generated using either lookup tables or developed using a subset of previously generated multiples. Usually, the entire multiplicand is multiplied by one multiplier digit to generate a partial product in each cycle. The partial product is added to an intermediate product register that holds the previously accumulated partial products. Several existing designs for decimal multiplication generate and store multiples of the multiplicand before partial product generation. Then the multiplier digits are used to select the appropriate multiple as the partial product [2]. The multiplier presented in [4] makes use of a secondary set of multiples generated using combinational logic. This design uses dedicated hardware operating at high frequencies with relatively low latencies. The multiplier in [9] stores intermediate product digits in a less restrictive, redundant format called the overloaded decimal representation that reduces the delay of the iterative portion of the multiplier. An alternative approach is to generate the partial product as needed. This leads to less wiring and elimination of registers to store multiples of the multiplicand [5]. An integral building block of a decimal digit by digit multiplier is the single digit multiplier. The single digit multiplier in [8] uses a standard 4×4 unsigned binary multiplier that generates an 8-bit binary output that needs to be corrected to two decimal digits.

Decimal floating point has 3 representations – 32 bit format with 7 significand digits, 64 bit format with 16 significand digits and 128 bit format with 34 significand digits. Fixed point multiplication of significand digits is an integral component of floating point multiplication. This paper presents a novel design for double digit fixed-point decimal multiplication that offers low latency and high throughput. The proposed multiplier performs 2 digit multiplications simultaneously in one clock cycle. When multiplying two n-digit operands to produce a 2n-digit product, the design has latency of $\lceil (n/2)+1 \rceil$ cycles. DDDM for 7 digits, 16 digits and 34 digits are simulated using Leonardo Spectrum from Mentor Graphics Corporation using ASIC Library. FPGAs are increasingly being used for improving performance by scientific computing community to implement floating-point based hardware accelerators. Hence, this paper also presents area and delay comparison for implementations of 7 digits, 16 digits and 34 digits DDDM on different families of Xilinx, Altera, Actel and Quick logic FPGAs.

The organization of the paper is as follows: Initially, the approach for a double digit multiplication is discussed. Then DDDM for various lengths are synthesized using ASIC Library, and the results are tabulated. The designs are then compared with single digit decimal multipliers (SDDM) in terms of area and delay. Finally, the paper concludes by tabulating a comparison of the proposed design for various lengths on different families of FPGA.

II. DECIMAL MULTIPLICATION

A decimal multiplier multiplies an n-digit multiplicand, A, by an n-digit multiplier, B producing a $2n$-digit product, P. A straightforward approach to decimal multiplication is to iterate over the digits of the multiplier, B, and based on the value of the current digit, B_i, successively add multiples of A to a product register [12]. The multiplier is accessed from least significant digit to most significant digit, and the product register is shifted one digit to the right after an iteration corresponding to division by 10. This approach allows an n-digit adder to be used to add the multiples of A to the partial product register. The multiples 2A through 9A, called primary multiples, are calculated at the start of the algorithm and stored along with A to reduce delay. The disadvantages of this approach are the enormous area or delay required for generating all the eight multiples, and the eight additional registers needed to store these multiples. An alternative method is to find a reduced set of multiples called secondary multiples. For example, if 2A, 5A, and 8A are computed and stored along with A, all the other multiples can be obtained with, at most, a single addition. This reduced set of multiples is called a secondary set, as no more than two members of the set need to be added to generate a missing multiple. Another reduced set of multiples comprising A, 2A, 4A, and 8A has a one-to-one correspondence with the weighted bits of a BCD digit. The disadvantage is that certain missing multiples can be generated only by the addition of 3 multiples from the reduced set. For example, the generation of 7A requires the addition of three multiples: A, 2A, and 4A. Although the secondary multiple approach reduces the delay or area and register count, it introduces the overhead of potentially one more addition for each iteration. The multiplier design proposed by [4] uses decimal carry-save addition to reduce this overhead. It gives a decimal multiplication algorithm suitable for high-performance with short cycle times. Since the floating point multiplier may need to handle operands up to 34 decimal digits further improvements in latency are suggested in this research using a double digit decimal multiplication technique.

III. DOUBLE DIGIT DECIMAL MULTIPLICATION (DDDM)

The block diagram for the DDDM is shown in Fig. 1. The 'Secondary Multiple Generation Block' generates secondary multiples 2A, 4A and 5A of length (n+1) digits. This is a purely combinational block with a maximum delay of 6 gates [4].
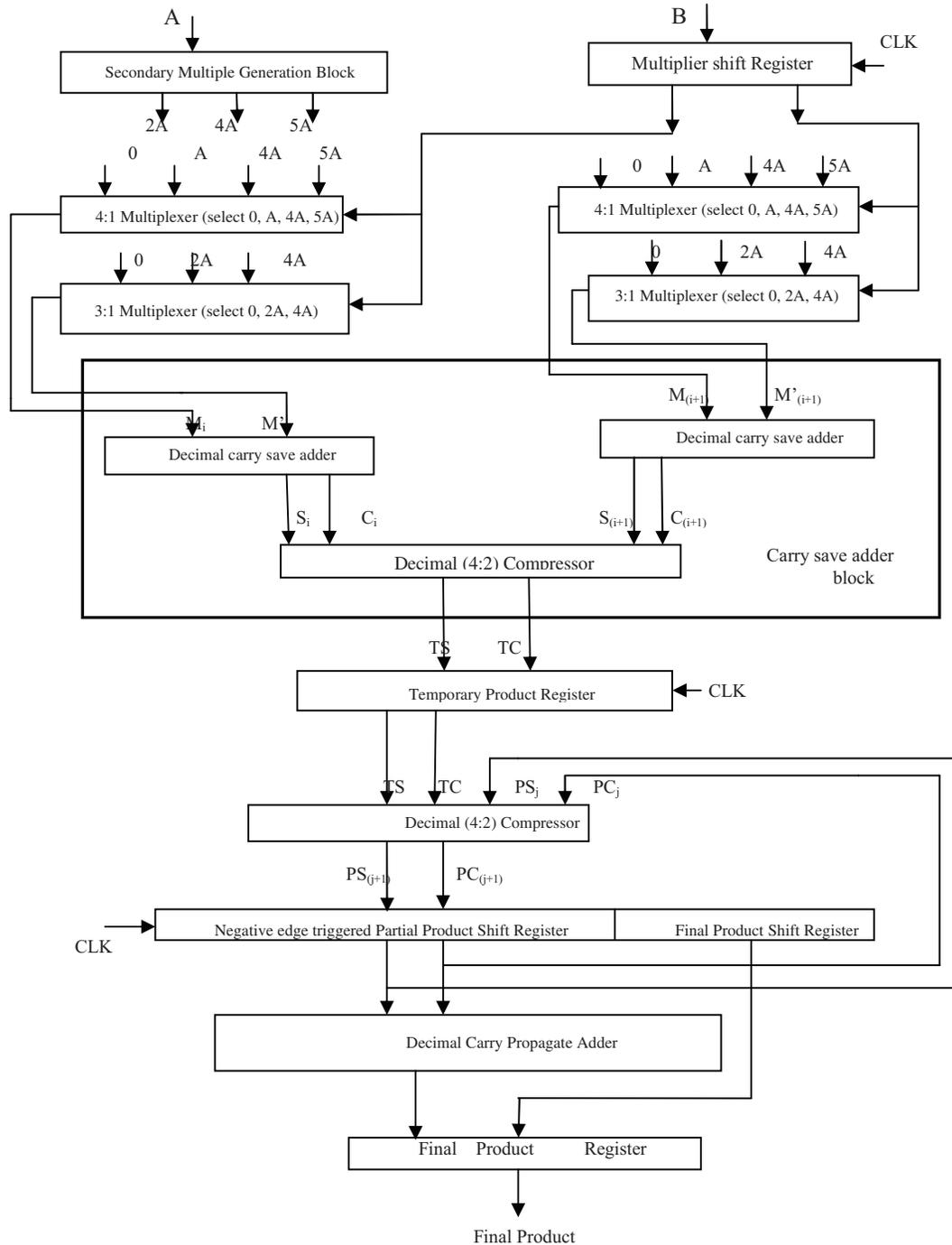
Figure 1. Block Diagram of the Double Digit Fixed Point Decimal Multiplication

The multiplier input, B is loaded into the 'Multiplier Shift Register' using an asynchronous load input. Suitable secondary multiples are selected by using two pairs of multiplexers for the two digit multiplier shift register output using Table 1. The 'Decimal Carry save Adder Block' adds the two selected secondary multiplies using carry save addition and generates an (n+1) digit sum output (S_i) and an (n+1) bit carry output (C_i). Similar addition is done by the second decimal carry save adder to produce $S_{(i+1)}$ and $C_{(i+1)}$. These 4 outputs are now added by a 4:2 compressor to give temporary sum (TS) and temporary carry (TC) values, of length (n+1) digits and (n+1) bits respectively. The TS and

TC values stored in the 'Temporary Product Registers' are added with the shifted output of the previous partial product (PS_i and PC_i) in the 'Partial Product Register' using a 4:2 compressor to get a new partial product. The 'Decimal Carry save Adder Block' adds the two selected secondary multiplies using carry save addition and generates an (n+1) digit sum output (S_i) and an (n+1) bit carry output (C_i). Similar addition is done by the second decimal carry save adder to produce $S_{(i+1)}$ and $C_{(i+1)}$. These 4 outputs are now added by a 4:2 compressor to give temporary sum (TS) and temporary carry (TC) values, of length (n+1) digits and (n+1) bits respectively.

TABLE 1: RECODING OF DIGITS OF B_i

B_i	M_i	M'_i	B_i	M_i	M'_i
0	0	0	5	5A	0
1	A	0	6	4A	2A
2	0	2A	7	5A	2A
3	A	2A	8	4A	4A
4	4A	0	9	5A	4A

The TS and TC values stored in the 'Temporary Product Registers' are added with the shifted output of the previous partial product (PS_i and PC_i) in the 'Partial Product Register' using a 4:2 compressor to get a new partial product. The last two digits of the partial product formed is a part of the final product. The new partial product is stored in the 'Partial Product Shift Register' at the negative edge of the clock in shifted form. For this purpose, the data in the 'Final Product Shift Register' is shifted for 2 digits during the previous positive edge, giving room to store the new 2 digits of the final product during the negative clock edge. For each iteration cycle the multiplicand is multiplied by 2 digits of the multiplier. The partial product formed is shifted by 2 digits and the process is repeated for $\lceil (n/2) \rceil$ iterations. After $\lceil (n/2) \rceil$ iterations the final product is in the form of carry save' sum and carry is available at the output of the 'Partial Product Shift Register'. This is then passed to a 'Decimal Carry Propagate Adder', which is actually a decimal incrementer. Decimal Incrementer shown in Fig. 2 is a circuit that adds a single bit to a decimal digit along with the carry in, and gives the result in decimal with a carry out (C_{oi}) as given in equation (1).

$$. C_{oi}= X_{i(3)} C_i C_{o(i-1)} + X_{i(3)} X_{i(0)} C_i + X_{i(3)} X_{i(0)} C_{o(i-1)} \text{ --(1)}$$

The C_{out} is generated after 2 gate delays for each digit. For 'n' digit multiplication, the ripple delay for C_{out} at the final Decimal Propagate Adder is 2n gate delays, the maximum complexity of the gate being 5 input AND gate. The Boolean expressions for a single digit Decimal Incrementer are given in equations 2-5. The total delay of the 'Decimal Carry Propagate Adder' is the delay of one digit Decimal Incrementer and 2n gate delays. This is much less than the delay of an 'n' digit BCD ripple adder. The adder output is then stored in the 'Final Product Register'. The final product is available after $\lceil (n/2)+1 \rceil$ cycles.
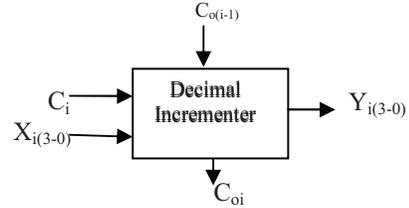

Figure 2. Decimal Incrementer

IV. SYNTHESIS RESULTS

The design was implemented in VHDL using Leonardo Spectrum from Mentor Graphics Corporation with ASIC Library for 7 digit, 16 digit and 34 digit multipliers. The simulations results are given in Tables 2, 3 and 4.

TABLE 2: AREA AND DELAY FOR VARIOUS BLOCKS OF DOUBLE DIGIT DECIMAL MULTIPLIER (7-DIGIT × 7-DIGIT)

Component (7 digit)	Area (μm^2)	Delay (ns)
Secondary multiple generation block (SMG)	1355	3.12
Multiplier shift register (MSR)	648	0.56
Multiplexer blocks (MUX)	1270	1.96
Carry save addition block (CSA)	4894	8.28
Temporary Product register (TPR)	387	0.44
Decimal 4:2 compressor (4:2C)	2442	5.45
Partial Product shift register (PPR)	913	0.79
Decimal carry propagate adder (CPA)	844	18.16

TABLE 3: AREA AND DELAY FOR VARIOUS BLOCKS OF DOUBLE DIGIT DECIMAL MULTIPLIER (16-DIGIT × 16-DIGIT)

Component (16 digit)	Area (μm^2)	Delay (ns)
Secondary multiple generation block (SMG)	3106	3.12
Multiplier shift register (MSR)	648	0.56
Multiplexer blocks (MUX)	2694	2.16
Carry save addition block (CSA)	10550	8.28
Temporary Product register (TPR)	774	0.44
Decimal 4:2 compressor(4:2C)	4883	5.45
Partial Product shift register (PPR)	1838	1.04
Decimal carry propagate adder (CPA)	1688	34.88

$$Y_{i(3)} = X_{i(3)}\overline{C_i}\overline{C_{0(i-1)}} + (X_{i(2)}X_{i(1)}X_{i(0)} + X_{i(3)}\overline{X_{i(0)}})(C_i \oplus C_{0(i-1)}) + \overline{X_{i(3)}}X_{i(2)}X_{i(1)}C_iC_{0(i-1)} - - - - - - - - - - -(2)$$

$$Y_{i(2)} = X_{i(2)}\overline{C_i}\overline{C_{0(i-1)}} + (X_{i(1)}X_{i(0)} + \overline{X_{i(3)}}X_{i(2)})(C_i \oplus C_{0(i-1)}) + (\overline{X_{i(3)}}X_{i(2)}\overline{X_{i(1)}} + \overline{X_{i(2)}}X_{i(1)})C_iC_{0(i-1)} - - - -(3)$$

$$Y_{i(1)} = X_{i(3)}\overline{C_i}\overline{C_{0(i-1)}} + (X_{i(1)}\overline{X_{i(0)}} + \overline{X_{i(3)}}X_{i(1)}X_{i(0)})(C_i \oplus C_{0(i-1)}) + \overline{X_{i(3)}}\overline{X_{i(1)}}C_iC_{0(i-1)} - - - - - - - - - - - - (4)$$

$$Y_{i(0)} = X_{i(0)} \oplus C_i \oplus C_{0(i-1)} -(5)$$

TABLE 4: AREA AND DELAY FOR VARIOUS STAGES OF
DOUBLE DIGIT DECIMAL MULTIPLIER (34-DIGIT × 34-DIGIT)

Component (34 digit)	Area (μm²)	Delay (ns)
Secondary multiple generation block (SMG)	6607	3.12
Multiplier shift register (MSR)	2664	0.73
Multiplexer blocks (MUX)	5564	2.18
Carry save addition block (CSA)	21417	8.28
Temporary Product register (TPR)	1548	0.44
Decimal 4:2 compressor (4:2C)	10045	5.45
Partial Product shift register (PPR)	3858	0.92
Decimal carry propagate adder (CPA)	3471	70.20

Fig. 3 and Fig. 4 show that the maximum area is occupied by the Carry Save Adder block, and maximum delay is for Decimal Carry Propagate Adder block for all implementations.

Table 5 indicates the comparison of area and delay parameters of DDDM with SDDM in [4]. It is noted that even though the area is increased by 50%, the speed to complete n-digit x n-digit multiplication is almost doubled as in the case of 16 digits ad 34 digits. But, the speed is only 1.44 times for 7 digits because the number of cycles required is $\lceil (n/2)+1 \rceil$ which is equal to 5 cycles, compared to a single digit multiplier that requires 8 cycles. Fig. 5 and Fig. 6 give a comparison of area and delay respectively of DDDM and SDDM for various lengths.
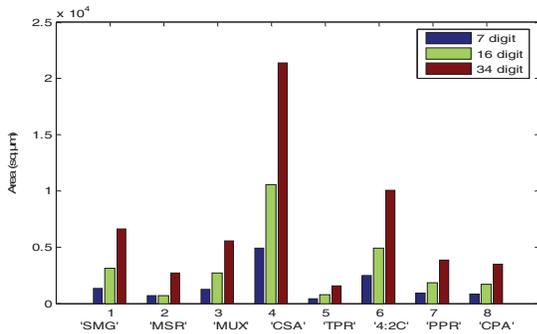

Figure 3: Area for various blocks of Double Digit Decimal Multipliers (7, 16, & 34 digits)

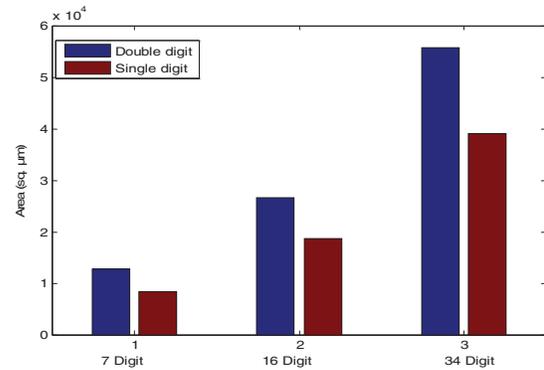Figure 4: Delay for various blocks of Double Digit Decimal Multipliers (7, 16, & 34 digits)

TABLE 5: COMPARISONS OF DDDM AND SDDM

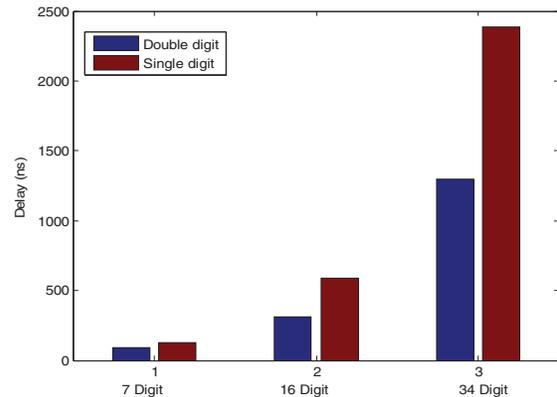| Digits	Parameters	Double digit	Design of [4]	Ratio
7-Digit	Area (μm²)	12713	8268	1.537
	Delay for 1 cycle (ns)	19.54	17.29	1.13
	Delay to complete 7 digit x 7 digit multiplication(ns)	90.8	130.4	0.696
16-Digit	Area (μm²)	26539	18717	1.42
	Delay for 1 cycle (ns)	38.10	36.24	1.05
	Delay to complete 16 digit x 16 digit multiplication(ns)	313.92	591.26	0.531
34-Digit	Area (μm²)	55714	39063	1.42
	Delay for 1 cycle (ns)	76.08	69.33	1.09
	Delay to complete 34 digit x 34 digit multiplication(ns)	1295.28	2388.4	0.542

Figure 5: Area comparison

Figure 6: Delay comparison

A. *Implementations in different FPGA families*

The designs were implemented in different families of various FPGAs and the results are tabulated. Table 6 shows the implementation results of DDDM for 7, 16 and 34 digits on various FPGAs. The study reveals that efficient mapping of the double digit multiplier is achieved when Xilinx FPGA devices are used.

TABLE 6: IMPLEMENTATION RESULTS OF DOUBLE DIGIT DECIMAL MULTIPLIERS ON FPGAs

FPGA	Utilized Area (Cells)			Maximum Frequency (MHz)			Delay (ns)		
	34 digit	16 digit	7 digit	34 digit	16 digit	7 digit	34 digit	16 digit	7 digit
ALTERA	12191	2405	1072	32.7	25.1	25.2	194.33	85.07	37.54
ACTEL	18759	3937	1819	14.7	14.6	15.1	346.58	129.88	65.86
XILINX	5765	1152	525	59.9	58.6	45.2	99.49	36.75	25.23
Quick Logic	12427	2500	1139	8.6	8.6	8.5	870.73	297.14	132.46

V.   CONCLUSION

This paper proposed double digit decimal fixed point multipliers that can be used in floating point multiplier circuits. It is noted that even though area is increased by 50%, the speed to complete n-digit × n-digit multiplication is almost doubled. This design leads to more regular VLSI implementation, and does not require special registers for storing easy multiples. The design was validated using lengths of 7 digits, 16 digits, and 34 digits multipliers that are required for all the three formats of floating point decimal multiplication. The synthesized design has a latency of 90.8 ns, 313.92 ns, 1295.28 ns respectively for 7 digits, 16 digits, and 34 digits fixed point multipliers. This design can be pipelined to increase the throughput. The latency for the multiplication of two n-digit BCD operands is $\lceil (n/2)+1 \rceil$ cycles, and a new multiplication can begin every '$n/2$'cycle. The design for 7, 16 and 34 digits of fixed point multipliers was also implemented on Xilinx, Altera, Actel and Quick logic FPGAs. Future research focuses on implementing floating point multipliers using the proposed fixed point multipliers.

REFERENCES

[1]   J. J. Bradley, B. L. Stoffers, T. R. S. Jr., and M. A. Widen, "Simplified Decimal Multiplication by Stripping Leading Zeros," U.S. Patent, Jun 1986. #4,615,016.

[2]   F. Y. Busaba, C. A. Krygowski, W. H. Li, E. M. Schwarz, and S. R. Carlough, "The IBM Z900 Decimal Arithmetic Unit," in Asilomar Conference on Signals, Systems, and Computers, vol. 2, pp. 1335–1339, November 2001

[3]   M. A. Erle, J. M. Linebarger, and M. J. Schulte, "Potential Speedup Using Decimal Floating-Point Hardware." 36th Asilomar Conference on Signals, Systems and Computers, Nov 2002.

[4]   M. A. Erle and M. J. Schulte, "Decimal Multiplication Via Carry-Save Addition," IEEE 14th International Conference on Application-specific Systems, Architectures and Processors, pp. 348-358, June 2003

[5]   M.A. Erle, E.M. Schwarz, M.J. Schulte, "Decimal multiplication with efficient partial product generation", 17th IEEE Symposium on Computer Arithmetic, 2005 ARITH-17 2005 pp. 21- 28, June 2005

[6]   R. L. Hoffman, T. L. Schardt, "Packed Decimal Multiply Algorithm" IBM Technical Disclosure Bulletin, vol18, pp.1562–1563,Oct'75

[7]   IEEE Standards Committee, "IEEE Standard for Floating-Point Arithmetic." http : // 754 r.ucbtest.org / drafts/754r.pdf, February 2003.

[8]   G. Jaberipur, A. Kaivani, "Binary-coded decimal digit multipliers", Computers & Digital Techniques, IET Vol. 1,Issue 4, July 2007, pp 377–381

[9]   R. D. Kenney, M. J. Schulte and M. A. Erle, "A High-Frequency Decimal Multiplier," IEEE 14th International IEEE international conference on Computer Design (ICCD'04), pp. 22-29, Oct 2004

[10]   R. H. Larson, "High Speed Multiply Using Four Input Carry Save Adder," IBM Technical Disclosure Bulletin, pp. 2053–2054, December 1973.

[11]   Ohtsuki, Y. Oshima, S. Ishikawa, K. Yabe, M. Fukuta, "Apparatus for Decimal Multiplication," U.S. Patent, Jun 1987,#4,677,583.

[12]   R. K. Richards, Arithmetic Operations in Digital Computers. New Jersey: D. Van Nostrand Company, Inc., 1955.