

高速低功率乘-累加器微架構與電路設計

學生：林書玄

指導教授：黃 威

國立交通大學電子工程學系電子研究所碩士班

摘 要

本論文針對處理器資料路徑提出一個在電路階層功率與速度最佳化的方法。利用高效率的乘法演算法，一個高速的乘-累加器微架構在本論文中被實現。根據此一高速的乘-累加器微架構設計，一個基於電晶體層次之高速低功率的乘-累加器也被實現。電晶體尺寸大小、電源電壓、以及臨界電壓作為調整參數，聯合這些參數採用本論文提出功率與速度最佳化的方法，可以使得動態的功率消耗減少原來的一半並使得速度提昇百分之二十。降低漏電流的設計方法於第四章被討論，針對在微架構階層一個資料路徑的功率與速度最佳化方法在第五章被討論。

本論文以 TSMC 0.13 μm CMOS 技術實現。一個高速低功率 16X16+32 的乘-累加器利用微架構與電路的設計技巧在本論文中被實現，一個乘-累加的運算其最長路徑所需的時間在 2 奈秒內，動態功率消耗為 10 毫瓦。

High-Speed and Low-Power Multiplier-Accumulator Micro-Architecture and Circuit Design

Student : Shu-Hsuan Lin

Advisor : Dr. Wei Hwang

Department of Electronics Engineering & Institute of Electronics

National Chiao-Tung University

ABSTRACT

A power-speed optimization technique of circuit level for a datapath of processors is proposed in this thesis. By using efficient multiplication algorithms, a high-speed multiplier-accumulator micro-architecture is designed in this thesis. According to this high-speed micro-architecture design, a low-power transistor level multiplier-accumulator is also implemented. Take the transistor size, the supply voltage, and the threshold voltage as tuning variables which are optimized jointly in terms of power and speed in this thesis which can reduce the dynamic power to one half and can increase the speed to 20%. Design techniques of leakage current suppression are discussed in chapter 4. The micro-architecture optimization methods in terms of power and speed are also examined in chapter 5.

All the results are simulated in TSMC 0.13 μm CMOS technology. Making use of micro-architecture and circuit level design techniques, the critical path of a 16X16+32 multiplier-accumulator operation is within 2ns, the dynamic power consumption is below 10 mW.

Acknowledgements

I would like to thank my advisor, Prof. Wei Hwang. Prof. Wei Hwang has been an invaluable source of guidance, suggestions, and encouragement throughout this thesis.

I would also like to thank Chung-Hsien Hua, who is currently working toward the Ph.D degree at NCTU. His advantageous suggestions helped me a lot in my thesis research.

My friends at NCTU have been a valuable source of moral support. I would like to thank all of them for their friendship during my years at NCTU.

My friend Ivy, she is the most important source of inspiration and happiness in my years at NCTU.

My family, who have always been a tremendous source of encouragement and confidence.



Table of Contents

Chapter 1	Introduction.....	1
Chapter 2	Power-Speed Tradeoffs in Datapath Structures.....	3
2.1	Overview of Multiplier-Accumulator Unit.....	3
2.2	Logical Effort.....	4
2.2.1	Model of Logic Gate	4
2.2.2	Multistage Logic Networks	6
2.2.3	Calculating Logical Effort and Calibrating the Model	9
2.3	Low Power Techniques in Datapath Structures.....	14
2.3.1	Power Dissipation in CMOS Circuits	14
2.3.2	Power Minimization Techniques	14
2.3.3	Design Time Power Reduction Techniques	15
2.3.4	Run Time Power Management	17
2.3.5	Reducing the Power in Sleep Mode.....	18
2.4	Theoretical Studies	19
2.4.1	Concepts of Methods for True Power Minimization.....	19
2.4.2	Energy-Delay Sensitivities.....	20
2.4.3	Circuit-Level Optimization.....	25
2.4.4	Micro-Architectural Optimization	26
2.5	Conclusions.....	27
Chapter 3	High-Speed Multiplier-Accumulator Micro-Architecture Design.....	28
3.1	Background.....	28
3.1.1	Basic Concepts.....	28
3.1.2	Booth Algorithm.....	30
3.1.3	Partial Product Matrix Topologies	31
3.1.4	Final Addition.....	37
3.2	Booth Recoding Schemes.....	37
3.2.1	Standard Scheme-Five Signals.....	38
3.2.2	Race Free Scheme-Four Signals	38
3.2.3	PR3 Scheme-Three Signals.....	38
3.2.4	Sign-Select Scheme-Three Signals	38
3.2.5	Experimental Evaluations.....	43
3.3	Partial Product Matrix Topologies	44
3.3.1	Carry Save Adder Array Topology.....	44
3.3.2	Column Compression Tree Topology.....	45

3.4	High Performance Parallel Adders.....	53
3.4.1	Design Strategies	53
3.4.2	Several High Performance Adders.....	53
3.5	Synthesized Results	60
3.6	Conclusions.....	63
Chapter 4	Power-Speed Optimization in Multiplier-Accumulator Circuit Design.....	65
4.1	Primitive Gates Calibration	65
4.2	Low Power Booth Recoder Design.....	65
4.2.1	Booth Selector Design.....	67
4.2.2	Booth Encoder Design.....	69
4.3	Power-Speed Optimization of MAC.....	73
4.3.1	Evaluation of XOR gate	73
4.3.2	The 5-2 Compressor's Family.....	76
4.3.3	Proposed Practical Power-Speed Optimization Procedure	81
4.3.4	Power-Speed Optimization of Column Compression Stage	81
4.3.5	Power-Speed Optimization of Final Adder	86
4.4	Managing Standby and Active Mode Leakage Power	89
4.4.1	Leakage Components	89
4.4.2	Standby Mode Leakage Control	90
4.4.3	Active Mode Leakage Control	91
4.5	Conclusions.....	94
Chapter 5	Micro-Architecture Level Optimization Related to Multiplier-Accumulator	95
5.1	Common DSP Processors Architecture	95
5.2	MAC Units for Common DSP Processors.....	97
5.3	Power-Optimum Pipelining.....	100
5.3.1	Pipelining versus Supply Voltage	100
5.3.2	Optimum Logic Depth per Pipeline Stage	101
5.4	Parallelism Exploitation to Improve Performance	102
5.5	Reconfigurable Power-Aware Architecture Design	106
5.5.1	Variable Precision Multiplier Architecture.....	106
5.5.2	Power Aware Variable Pipeline Stage Architecture	109
5.6	Conclusions.....	111
Chapter 6	Conclusions.....	112
6.1	Summary	112
6.2	Future Work	112

References 114



List of Tables

Table 2.1	Best number of stages to use for various path efforts.....	8
Table 2.2	Power-speed optimization techniques.....	15
Table 3.1	Modified Booth algorithm.....	31
Table 3.2	Selected high speed recoding scheme.....	38
Table 3.3	Standard encoding scheme.....	39
Table 3.4	Race free II scheme.....	40
Table 3.5	PR3 encoding scheme.....	41
Table 3.6	Sign select scheme.....	42
Table 3.7	Comparison of Booth Recoder without any synthesized constrains.....	44
Table 3.8	Comparison of Booth Recoder with 100ps timing constrains.....	44
Table 3.9	Synthesized result of CSA array with and without timing constrains.....	45
Table 3.10	Synthesized result of CCT topology with and without timing constrains.....	49
Table 3.11	Synthesized results of 32-bit adder without timing constrain.....	54
Table 3.12	Synthesized results of 32-bit adder with 0.5ns timing constrain.....	54
Table 3.13	Synthesized result of 16X16+32 MAC macro without timing constrain.....	61
Table 3.14	Synthesized result of 16X16+32 MAC macro with 2ns timing constrain.....	61
Table 3.15	Synthesized result of 16X16+40 MAC macro without timing constrain.....	62
Table 3.16	Synthesized result of 16X16+40 MAC macro with 1ns timing constrain.....	62
Table 4.1	Comparisons of different 5-2 compressors.....	77
Table 4.2	Numerical results of different implementations of 5-2 compressors.....	81
Table 5.1	MAC operations of C64X DSP processor.....	104
Table 5.2	Control words supported by the configurable multiplier.....	108

List of Figures

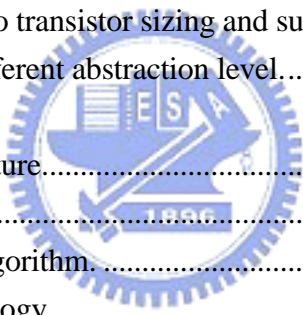
Fig. 2.1	Conceptual model of a CMOS logic circuit.	5
Fig. 2.2	Delay as a function of electrical effort.....	7
Fig. 2.3	Calculating logical effort of a simple gate.....	11
Fig. 2.4	Simulated delay of inverters driving various loads.	12
Fig. 2.5-1	Test circuit of inverter calibration for electrical effort $h=0$	12
Fig. 2.5-2	Test circuit of inverter calibration for electrical effort $h=2$	13
Fig. 2.5-3	Test circuit of inverter calibration for electrical effort $h=4$	13
Fig. 2.6	Static current from V_{DDL} to V_{DDH}	17
Fig. 2.7	A dynamic voltage-scaled system.....	18
Fig. 2.8	Sleep transistors and MTCMOS scheme.....	19
Fig. 2.9	Energy-speed optimization process.	21
Fig. 2.10	A sketch map of sensitivity to supply voltage.....	24
Fig. 2.11	A sketch map of sensitivity to threshold voltage.....	24
Fig. 2.12	Energy-delay sensitivities of V_{dd} , W , and V_{th}	25
Fig. 2.13	Joint optimization to transistor sizing and supply voltage.	25
Fig. 2.14	Optimization of different abstraction level.....	27
		
Fig. 3.1	Parallel MAC structure.....	29
Fig. 3.2	Dot diagram.	29
Fig. 3.3	Modified Booth Algorithm.	31
Fig. 3.4 (a)	Single array topology	32
Fig. 3.4 (b)	Double array topology	33
Fig. 3.4 (c)	Higher order array topology	34
Fig. 3.5	[4:2] Compressor.	33
Fig. 3.6	Three dimensional structural flatten into two dimensional plane.....	35
Fig. 3.7	Wallace's strategy v.s. Dadda's strategy.	36
Fig. 3.8	Wallace's and Dadda's strategy in 6X7 multiplication.	36
Fig. 3.9	Standard encoder and decoder.	39
Fig. 3.10	Race free encoder and decoder.	40
Fig. 3.11	PR3 encoder and decoder.....	41
Fig. 3.12	Sign select encoder and decoder.	42
Fig. 3.13	Booth recoding macro.	43
Fig. 3.14	16X16+32 MAC using carry save array.....	47
Fig. 3.15	Logical decomposition of [4:2] and [5:2] compressor.	46
Fig. 3.16	Dot diagram of a 16X16 2's complement MAC with arbitrary accumulator.	50
Fig. 3.17	A modified 16X16 2's complement MAC with arbitrary accumulator.	51

Fig. 3.18	Bit slice for 10:2 reduction.....	49
Fig. 3.19	The reduction process by using column compression tree topology.....	52
Fig. 3.20	Signal arrival profile of CSA array topology.....	55
Fig. 3.21	Signal arrival profile of CCT topology.....	55
Fig. 3.22	32-bit Kogge-Stone prefix adder.....	56
Fig. 3.23	32-bit Han-Carlson prefix adder.....	57
Fig. 3.24-1	32-bit Conditional Carry Adder.....	58
Fig. 3.24-2	32-bit Conditional Carry Adder.....	59
Fig. 3.25	A 16X16+40 MAC from Synopsys Design Ware Libraries.....	61
Fig. 3.26	Normalized result from Table 3.12.....	63
Fig. 3.27	Normalized result from Table 3.14.....	63
Fig. 3.28	Normalized result from Table 3.16.....	64
Fig. 4.1	Simulated delay of NAND2 driving various loads.....	66
Fig. 4.2	One bit Booth selector.....	67
Fig. 4.3	Test circuit and results of deciding the size of T-G and inverter.....	68
Fig. 4.4	Encoder delay estimation and simulation.....	69
Fig. 4.5	The critical path of Booth recoder.....	70
Fig. 4.6 (a)	Power distribution of Booth recoder under miniumu delay.....	70
Fig. 4.6 (b)	Trade speed for power reduction.....	71
Fig. 4.6 (c)	Power distribution and profile of Booth recoder.....	72
Fig. 4.6 (d)	Power distribution and profile of Booth recoder.....	72
Fig. 4.7	Six different designs of XOR gates.....	74
Fig. 4.8	Evaluation of XOR gates using logical effort.....	75
Fig. 4.9 (a)	The 5-2 compressor's family: 3-2 adder.....	76
Fig. 4.9 (b)	The 5-2 compressor's family: 4-2 compressor.....	76
Fig. 4.9 (c)	The 5-2 compressor's family: 5-2 compressor.....	77
Fig. 4.10-1	Decompositions of the 5-2 compressors: module1.....	79
Fig. 4.10-2	Decompositions of the 5-2 compressors: module2 and module3.....	78
Fig. 4.11	High-speed and low-power computational kernel of compressors.....	80
Fig. 4.12	Proposed powe-speed optimizatiom procedure.....	82
Fig. 4.13	Critical path and non-critical path in the column compressor stage.....	82
Fig. 4.14	Power-speed optimization via transistor sizing.....	83
Fig. 4.15	Power-speed optimization via threshold voltage scaling.....	84
Fig. 4.16	Power-speed optimization via supply voltage scaling.....	85
Fig. 4.17 (a)	Individual optimization of column compression stage.....	87
Fig. 4.17 (b)	Joint optimization of column compression stage.....	87
Fig. 4.18 (a)	Individual optimization of K-S adder.....	88
Fig. 4.18 (b)	Joint optimization of K-S adder.....	88

Fig. 4.19	MOSFET leakage components.	90
Fig. 4.20	Staking effect and VT CMOS.	91
Fig. 4.21	Schematic of dynamic V_t (DVTS) scaling.	93
Fig. 5.1	Harvard architecture.	96
Fig. 5.2	Skeleton of a MAC unit in DSP processors.	98
Fig. 5.3	Block diagram of the MAC functional unit.	99
Fig. 5.4	Micro-architectural design options.	101
Fig. 5.5	Parallelism exploitation in DSP processors.	103
Fig. 5.6	Block diagram of C64X DSP core.	104
Fig. 5.7	One combination of 16-bit SIMD instruction	105
Fig. 5.8	Four combinations of 16-bit SIMD instruction.	106
Fig. 5.9	Variable precision multiplier architecture.	107
Fig. 5.10	Recursive variable precision multiplier architecture.	110
Fig. 5.11	Reconfigurable 4-stage pipeline.	111

