

Contents

List of Figures	III
List of Tables	V
1 Introduction	1
2 Cache Review and Modeling Power Consumption	5
2.1 Cache Review	6
2.1.1 Review of Cache Organization	6
2.1.2 Review of Cache Operation	8
2.1.2.1 Cache Read Access	8
2.1.2.2 Cache Refill Access	9
2.1.2.3 Cache Invalidate	9
2.2 Development of Cache Power Consumption Model	9
2.2.1 Characterization of Energy Consumption Model	10
2.2.1.1 Decoding Path	11
2.2.1.2 Wordline	15
2.2.1.3 Bitline	15
2.2.1.4 Input / Output Lines	16
2.2.2 Energy Consumption Model in Direct-mapped Cache	16
2.2.3 Energy Dissipation in T0 I-cache	18
3 Study of Previous Work on Low Power Cache Design	21
3.1 Sub-banking	22
3.1.1 Experimental Study	22
3.1.2 Experimental Results	24
3.2 Reducing the Frequency of Tag Compares	30
3.2.1 Experimental Study	30
3.2.1.1 Branch and Jump	31
3.2.1.2 Interblock Sequential Flow	31
3.2.1.3 Frequency of Tag Compares	33
3.2.2 Experimental Methods	33
3.2.2.1 Evaluation Environment	33
3.2.2.2 Experimental Results	34

4	Background and Related Work on Using Gated Wordlines to Reduce the Number of Bitline Swings	39
4.1	Overview of Gated Wordline Technique to Reduce the Number of Bitline Swings	40
4.2	Background of Using Gated Wordline Technique	42
4.2.1	The 2-size Approach	43
4.2.1.1	The I-cache Refill	44
4.2.1.2	The I-cache Read Access	45
4.2.2	The 3-size Approach	46
4.2.2.1	The I-cache Refill	47
4.2.2.2	The I-cache Read Access	49
4.3	Review of MIPS Instruction Set	50
4.3.1	Type of Instructions	50
4.3.2	Instruction Formats	50
4.3.2.1	R-type Instructions	52
4.3.2.2	I-type Instructions	54
4.3.2.3	J-type Instructions	55
4.4	Previous Study on Code Compression	56
4.4.1	Approaches for Code Size Reduction	56
4.4.1.1	Short Instruction Encodings	57
4.4.1.1.1	MIPS16	57
4.4.1.2	Code Compression Using the Dictionary	59
5	Experimental Study on Using Gated Wordlines to Reduce the Number of Bitline Swings	61
5.1	Experimental Study	62
5.1.1	The 2-size Approach	62
5.1.1.1	Compression Technique for Medium Size	62
5.1.2	The 3-size Approach	64
5.1.2.1	Compression Technique for Short Size	65
5.2	Experimental Methods	71
5.2.1	Simulation	71
5.2.2	Experimental Result	79
5.2.3	Conclusion	83
6	Conclusion	85
	Bibliography	87

List of Figures

2.1	Logical organization of cache	6
2.2	Cache structure	7
2.3	Decoder circuit of T0 I-cache	11
2.4	A 2-to-4 predecoder block	11
2.5	Wordline drive logic	14
2.6	SRAM cell	14
2.7	Bitline precharged circuit	15
2.8	Breakdown of switching energy in T0 I-cache	18
3.1	Energy dissipation in option 1 where the whole cache line is read out from the data array	25
3.2	Energy dissipation in option 2 where sub-banking is used and only one instruction is read out from the data array	26
3.3	Energy dissipation in option 3 where cache organization is identical to option 2 but there is no tag compare	27
3.4	Energy saved when using sub-banking (comparing option 2 with option 1)	28
3.5	Energy saved when there is no tag compare as compared to when there is (comparing option 3 with option 2)	29
3.6	Percentage of tag compares in benchmark programs	35
3.7	Percentage of unnecessary tag compares that has been avoided in the benchmarks	36
4.1	CPU layout with compression and decompression blocks	40
4.2	Format layout of 2-size instructions	43
4.3	Format layout of 3-size instructions	43
4.4	Circuit to control the size of the instruction being 'written into' or 'read out from' SRAM array in the 2-size method	44
4.5	Style1: Circuit to control the size of the instruction being 'written into' or 'read out from' SRAM array in the 3-size method. In this style, two levels of metal are required.	46
4.6	Style 2: Circuit to control the size of the instruction being 'written into' or 'read out from' SRAM array in the 3-size method. In this style, only one level of metal is required.	46

4.7	MIPS RISC instruction formats	51
4.8	MIPS-II instruction encodings of integer subset	52
4.9	MIPS 16 Decompression	58
4.10	CPU layout with Dictionary for expansion of instructions	60
5.1	CPU instruction encodings of integer subset for the 2-size method	63
5.2	CPU instruction encodings of integer subset for the 3-size method	66
5.3	Variable-length immediate of medium-size instruction in the 2-size approach	72
5.4	Variable-length immediate of medium-size instruction in the 3-size approach	72
5.5	Dynamic compression ratio of various benchmark programs using the 2-size approach	73
5.6	Dynamic compression ratio of various benchmark programs using the 3-size approach	74
5.7	Percentage of reduction (saving) in bits read out in the 2-size approach	75
5.8	Percentage of reduction (saving) in bits read out in the 3-size approach	76
5.9	Instruction composition breakdown into medium and long instructions for the 2-size approach	77
5.10	Instruction composition breakdown into short, medium and long instructions for the 3-size approach	78
5.11	Summary of dynamic compression ratio in the benchmark programs when using 23-bit length for medium-size instruction	81
5.12	Summary of the percentage of reduction (saving) in bits read out in the benchmark programs when using 23-bit length for medium-size instruction	82

List of Tables

3.1	Benchmark programs	33
3.2	Statistics of branch and jump instructions in benchmark programs	34
4.1	Subfield definition	51
4.2	Sub-groups of R-type instructions	53
4.3	Sub-groups of I-type instructions	54
4.4	Immediates extension of the I-type instructions	55
4.5	J-type instructions	55
4.6	Target Extension of J-type instructions	56
5.1	Compression condition for I1_1	68
5.2	Compression condition for I2_1	69
5.3	Compression condition for I3	69
5.4	Compression condition for I2_1, I2_2, and I2_3 instructions	70

Acknowledgment

Thank you... Thank you... Thank you...

This thesis would not exist without many individuals.

Thank you my thesis supervisor, Krste Asanovic, for guiding and supporting me from the beginning through the end of this research with intellectual advice and insightful suggestions. This thesis would not have been as it is today without him.

Thank you Murali for advice, encouragement, reducing my stress level with fun talk every night, and the most important thing... being the editor in chief. This thesis would not have been completed without him.

Thank you Jessica Tseng and Seongmoo Heo for technical discussions and ideas for my thesis.

Thank you my sister, Taew, for listening to my numerous complaints during the stressful time.

Thank you P'Tan for always calling to check that I was still alive in the lab during my thesis crunching weeks.

Thank you Ant for late night's snack at 3 am.

Thank you all of my friends in MIT during undergrad and grad years. My life at MIT would have been so miserable without all of you.

Thank you Smooth Jazz 96.9 for keeping me awake for so many sleepless nights.

Finally, thank you my beloved parents for always making things in my life possible for me.