Arithmetic Data Value Speculation

A thesis submitted to

the School of Electrical and Electronic Engineering

of the University of Adelaide

by

Daniel R. Kelly

B.E. (Computer Systems)(Hons),

B.Sc. (Maths & Computer Science)

in fulfillment of the requirements for

the degree of Doctor of Philosophy (Microelectronics)

Commenced, April 2005

Thesis Submitted, August 2010

Declaration of Originality

This work contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of the thesis, when deposited in the University Library, being available for loan, photocopying, and dissemination through the library digital thesis collection, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue, the Australasian Digital Thesis Program (ADTP) and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Daniel R. Kelly
School of Electrical and Electronic Engineering
The University of Adelaide
September 13, 2010

Acknowledgements

I wish to acknowledge my principal supervisor, Dr. Braden J. Phillips, for all of his help and guidance in all the things that mattered, including: identifying a stimulating research project; establishing my scholarship that made the project possible; being flexible enough to allow my sometimes backwards approach but firm enough to reign me in where needed; and an attention to detail and conviction in the importance of good, clear writing (if in coming pages you disagree, I am happy to award him blame or credit ;P).

My co-supervisor Dr. Said Al-Sarawi has been invaluable in providing a very focused, objective critique—with an attention to detail and response time that was amazing. I am grateful for the time and resources that Said has committed to me when needed, above and beyond the call of duty.

Dr. Andrew Beaumont-Smith has been extremely generous in providing an industry perspective and finding opportunities to work in the job of a lifetime. Andrew's influence on me has been to contextualise my research and to change my perspective in a way that no one else has. *"Just get your PhD ASAP"*.

Fiona Tselentis has been behind the scenes gently pushing me along and helping in so many ways that she doesn't even know about.

Finally I would like to thank my parents for their support and encouragement. This thesis would not have been possible without them.

This project has been a rewarding and challenging undertaking. I will never, ever do one again. Without further ado, I present my thesis, warts and all.

Daniel R. Kelly

September 13, 2010

Abstract

Arithmetic approximation is used to decrease the latency of an arithmetic circuit by shortening the critical path delay or the sampling period so that result is not guaranteed to be correct for every input combination. Thus, an acceptable compromise between the circuit latency and the average probability of correctness drives the circuit design. Two methods of arithmetic approximation are:

temporally incompleteness where circuits quote the result before the critical path delay (overclocking); and

logically incompleteness where circuits use simplified logic, so that most input cases are calculated correctly, but the slowest cases are calculated incorrectly.

Arithmetic data value speculation (*ADVS*) is a speculation scheme based on arithmetic approximation, and is used to increase the throughput of a general purpose processor. *ADVS* is similar to branch prediction, an arithmetic instruction is issued to an exact arithmetic unit and an approximate arithmetic unit which provides an approximate result faster than the exact counterpart. The approximate result is forwarded to dependent operations so they may be speculatively issued. When the exact result is eventually known, it is compared to the approximate result, and the pipeline is flushed if they differ.

This thesis, *"Arithmetic Data Value Speculation"*, presents work in the field of digital arithmetic and computer architecture. A summary of current probabilistic arithmetic methods from the literature is provided, and novel designs of approximate integer arithmetic units are presented, including results from logical synthesis. A case study demonstrates approximate arithmetic units used to increase the average throughput of benchmark programs by speculatively issuing dependent operations in a RISC processor.

The average correctness of the approximate arithmetic units are shown to be highly data dependent, results vary depending on the benchmarks being run. In addition, the

average correctness when running benchmarks is consistently higher than for random inputs. Simulations show that many arithmetic operations are often repeated in the same benchmark, leading to a high variation in correctness. Speculative gains from one operation can be offset by speculation losses due repeated incorrect approximation of another approximate unit, so typical throughput gains through speculation in a general purpose processor pipeline are low. The minimum threshold correctness of an approximate arithmetic unit used for speculation is shown to be approximately 95%. Logic synthesis is used to determine power, area and timing information for approximate units implemented from novel algorithms, and show a reduction in arithmetic cycle latency for integer operations, and the expense of 50 % leakage and area, and 90 % dynamic power.

Value speculation can be complemented by result caching; repeated pipeline flushes can be avoided if the correct result is know before speculation, the average operation latency can be reduced, and caching can be used for operations that are difficult to approximate.

Publications

The following is a list of publications published or submitted during the Ph.D. candidature by the author.

Published book chapters

1. Daniel R. Kelly and Braden J. Phillips, "Arithmetic data value speculation", in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3740 NCS, pp. 353–366, 2005.

2. Daniel R. Kelly, Braden J. Phillips and Said Al-Sarawi, *Algorithm-ArchitectureMatching for Signal and Image Processing*, volume 73 of *Lecture Notes in Electrical Engineering*, chapter 4 "Operators for embedded systems: Approximate Multiplication and Division for Arithmetic Data Value Speculation in a RISC Processor". Springer, 1st edition, December 2010. ISBN: 978-90-481-9964-8.

Published conference papers

3. Daniel R. Kelly, Braden J. Phillips and Said Al-Sarawi, "An open source synthesisable model in VHDL of a 64-bit MIPS-based processor", in *Proceedings of SPIE—The International Society for Optical Engineering*, vol. 6414, (Adelaide, Australia), 2007.

4. Braden J. Phillips, Daniel R. Kelly and Brian W. Ng, "Estimating adders for a low density parity check decoder", in *Proceedings of SPIE—The International Society for Optical Engineering*, vol. 6313, (San Diego, CA, United States), 2006.

5. Braden J. Phillips, Cain D. Schmidt and Daniel R. Kelly, "Recovering data from USB flash memory sticks that have been damaged or electronically erased", in *e-Forensics '08: Proceedings of the 1st international conference on Forensic applications and techniques in telecommunications, information, and multimedia and workshop*, (ICST, Brussels, Belgium, Belgium), pp. 1–6, ICST (Institute for Computer Sciences, Social Informatics and Telecommunications Engineering), 2008.

6. Daniel R. Kelly, Braden J. Phillips and Said Al-Sarawi, "Approximate unsigned binary integer dividers for arithmetic data value speculation". In *Proceedings of the Conference on Design and Architectures for Signal and Image Processing (DASIP)*, Sophia Antipolis, France, September 22–24, 2009. http://www.ecsi-association.org/ecsi/dasip/dasip09.

7. Daniel R. Kelly, Braden J. Phillips and Said Al-Sarawi, "Approximate signed binary integer multipliers for arithmetic data value speculation". In *Proceedings of the Conference on Design and Architectures for Signal and Image Processing (DASIP)*, Sophia Antipolis, France, September 22–24, 2009. http://www.ecsi-association.org/ecsi/dasip/dasip09.

8. Daniel R. Kelly, Braden J. Phillips and Said Al-Sarawi, "Increasing throughput of a RISC system using arithmetic data value speculation". In *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, USA, November 1–4, 2009.

Contents

List of Figures

List of figures

List of Tables

List of Tables

List of Algorithms

List of Acronyms

ADPCM	adaptive differential pulse code modulation
AWCCL	average worst-case carry length, in bits
AWGN	additive white gaussian noise
CCITT	International Telegraph and Telephone Consultative Committee
CMOS	complementary metal oxide semiconductor
DFT	discrete fourier transform
FP	floating point
EPIC	Efficient Pyramid Image Coder
FA	full adder
FFT	fast fourier transform—an efficient algorithm to compute the discrete fourier transform (DFT) and its inverse.
GSM	Global System for Mobile communications: originally from Groupe Spécial Mobile
HA	half adder
IEEE	Institute of Electrical and Electronic Engineers
INT	integer
ISA	instruction set architecture
JPEG	Joint Photographic Experts Group

List of Algorithms

LSB	least significant bit
MPEG	Moving Picture Experts Group
MSB	most significant bit
MUX	Multiplexor
SNR	signal-to-noise ratio
QPSK	quadrature phase-shift keying
PGP	Pretty Good Privacy
SPEC	Standard Performance Evaluation Corporation
SPHERE	SPeech HEader REsources
ULP	Unit of least precision

Nomenclature

Mathematical notation

$f(n) \in O(n)$	$f(n) \leq g(n) \cdot k$, f is bounded above by g (up to constant factor) asymptotically
$f(n) \in \Omega(n)$	$f(n) \geq g(n) \cdot k$, f is bounded below by g (up to constant factor) asymptotically
$f(n) \in \Omega(g(n))$	$g(n) \cdot k_1 \leq f(n) \leq g(n) \cdot k_2$, f is bounded both above and below by g asymptotically
$\lfloor x \rfloor$	floor(x), greatest integer $\leq x$
$\lceil x \rceil$	ceiling(x), least integer $\geq x$
P_M	probability of correctness for a multiplier
P_C	probability of correctness for a counter

Standard notation

N	Number of bits in an operand; word length
a_N	Average worst-case carry length for N bit addition
l	Length in bits of a carry-chain in an adder
PC	Program Counter register
SP	Stack Pointer register

Multiplication

n	number of input bits to a counter

Nomenclature

m	number of output bits from a counter

Division

z	dividend
d	divisor
q	exact quotient after division $q = z/d$
d_H	least binary power greater than or equal to d
d_L	greatest binary power less than or equal to d
$\tilde{d}$	approximate divisor
$\tilde{D}$	$\log_2(d)$, or the number of zeros after the MSB of d
q_i	quotient after round i
r_i	remainder after round i
e_i	error in q_i after round i
n	number of bits in the unsigned binary representation of z and d
k	number of bits in the quotient, q
t	maximum number of division rounds
f	number of fractional bits maintained in q_i
m	constant multiplicand in the division round: $m = d - \tilde{d}$
p_1	most significant non-zero partial product in the binary multiplication $r_i \times m$
p_2	after p_1, the next most significant partial product in the binary multiplication $r_i \times m$