Lec-01: Introduction and Overview

ECE-720 topic 4: Innovations in Processor Design

Mark Aagaard
2008t1

Today's Lecture

- Historical perspective on computers
- Administrivia

Historical Perspective

Ancient History

Abacus (300BCE – 1930)

- Originated before written numerals
- Used to record numbers, does not do actual computation
- First preserved "counting board"
  - Greece 300BCE
  - Sand, pebbles, or disks used to mark numbers
- Modern abacus
  - China 1200 AD
  - Wires and rings
  - Variety of different styles (Chinese, Japanese, Russian)

Ancient History

- Al-Kashi (1393 – 1449)
  - Muslim mathematician
  - Developed specialized computing devices for astronomy (e.g. computing time of a lunar eclipse)
  - Perhaps first use of a decimal fraction
  - Calculated 2π to sixteen decimal places

Goldstine, 1972, pp5

Adding Machines

Schickard (1592 – 1635)

- Designed, but did not build, machine to do addition, subtraction, multiplication, division
- Documentation lost until 1950s(?), because Schickard and family died in the Plague

Blaise Pascal (1623 – 1662)

- Machines to do addition/subtraction by counting (1642)

Ancient History

Francois Vieta (1540 – 1603)

- First use of letters to represent unknowns or general parameters in mathematical equations

John Napier (1550 – 1617)

- Invented logarithms: multiplication and division computed using conversion tables and addition/subtraction
- Possibly first use of decimal point in arithmetic

Slide rule

- 1620 forerunner by Edmund Gunter
- 1632 William Oughtred and 1630 Richard Delamain

Goldstine, 1972, pp46

Arithmetic Machines

Gottfried Leibniz (1646 – 1716)

- Leibniz wheel: add, sub, mul, div (166?)
- "It is unworthy of excellent men to lose hours like slaves in the labour of calculation which could be safely relegated to anyone else if machines were used."
- Initiated the field of formal logic: "A general method in which all truths of the reason would be reduced to a kind of calculation."

Difference Engines

Charles Babbage

- Mechanical system to calculate values for polynomials (e.g. $x^2 + x + 41$)

Pehr Georg Scheutz (1785 – 1873)

- Inspired by articles about Babbage's Difference Engine, constructed own Difference Engine
- Novel design, more reliable
- First arithmetic engine with automatic printing

9

American Census 1890

- Herman Hollerith developed puch-card tabulation system
- Punch cards were same size as dollar bill, so that could use same sorting and filing equipment as banks and treasuries
- Inspired extensive development of electronic accounting machinery

11

Analytical Engine (1830—1840)

- Invented by Charles Babbage
- Never completed
- First programmable computing machine
- Based on punch cards similar to those used in semi-automated weavers
- Ada Lovelace wrote "programs" for the Analytic Engine
  - First computer programmer
- Partially constructed by British museum in ????, believed to operate correctly

10

- Computer = "person whose job it was to perform mathematical computations"

12

Howard Zuse (late 1930s---1940s)

- Mechanical computers in Germany
- First implementation of floating-point arithmetic

13

Harvard Mark I

- Developed by Howard Aiken at Harvard University
- Programmable arithmetic computation system
- Electro-mechanical
- Used scrolls of paper, similar to player piano, to control sequence of arithmetic operations to be performed
- Series of Harvard Mark computers
  - Mark-II: relays (electrical switches)
  - Mark-III: vacuum tubes, separated data and instruction memory
  - Mark-IV: vacuum tubes

14

John Atanasoff (1940s)

- Iowa State University
- ABC = “Atanasoff Berry Computer”
- Use of binary representation for numbers
- ABC was never fully operational
- Probably influenced Mauchly’s design for the ENIAC

15

ENIAC (approx 1944)

- ENIAC = “Electronic Numerical Integrator and Calculator”
- J. Presper Eckert and John Mauchly
- University of Pennsylvania
- 20 ten-digit registers
- 18,000 vacuum tubes
- 200 microseconds/addition
- First fully electronic computing system
- Approximately 1000 times faster than Mark I
- “The ENIAC is so fast, that to make use of six ENIACs would require all of the mathematicians to devote themselves exclusively to the task of finding problems for the computers to solve.”
- Programming was done by connecting wires and flipping switches

16

EDVAC memo (1944)

- EDVAC = “Electronic Discrete Variable Automatic Computer”
- Memo written by John von Neumann, Eckert, and Mauchly (but Eckert and Mauchly’s names were ommitted when memo distributed by Herman Goldstine)
- Key ideas
  - Program stored as sequence of numbers in memory
  - Program counter
  - Computation modules
- Patent dispute with UPenn crippled the project
- Finally operational in 1952

Manchester Mark-I / Baby (1948)

- Electro-mechanical
- First operational stored-program machine

MIT Whirlwind (1947---????)

- Application domain: real-time signal processing for radar
- Innovation: magnetic core memory
  - First reliable and inexpensive form of memory
  - Most common form of memory until 1970s
- 2048 16-bit words of memory

BINAC (1949) and UNIVAC (1951)

BINAC

- Ekert-Mauchly Computer Corporation (formed 1947)
- BINAC was first product of E-M Computer Corp
- E-M bought by Remington-Rand
- Remington-Rand renamed to Sperry-Rand

UNIVAC-I

- Product of Sperry-Rand
- First successful commercial computer
- 48 systems were built
- Price: US$ 250,000

EDSAC (1949)

- EDSAC= “Electronic Delay Storage Automatic Calculator”
- Maurice Wilkes visited ENIAC project at UPenn (1946)
- First full-scale operational stored program electronic computer
- Instruction set was based on accumulator

21

IAS Computer (19??)

- IAS = Institute for Advanced Studies (Princeton Univ)
- 1024 40-bit words
- 10 times faster than ENIAC

22

IBM 701 (1952)

- IBM’s first computer
- Based on ideas in IAS
- Sold 19 units

23

IBM Stretch (1955 – 1964)

- Design started in 1955
- Goals:
  - Performance: 100 times faster than existing computers
  - Development time: 5 years (Delivery in 1960)
  - Price: approx US$5M
- Reality:
  - 16 times faster than existing computers
  - Development time: 6 years
  - Price: US$13.5M
- Fate
  - Fastest computer in the world: 1961—1964
  - Selling price reduced to US$7.8M
  - Immediately discontinued

24

IBM Stretch

Why do we care today?

- Chief architects very influential (Blaaw, Brooks, Cocke)
- Technological influences still felt today
  - Architectural simulator and benchmark programs
  - Predecoding
  - Pipelining
  - Out-of-order execution (lookahead unit)
  - Speculative execution
  - Branch mispredict recovery
  - Precise interrupts
  - Fused multiply/add instruction

25

IBM Stretch

- Reasons for slowness
  - Clock speed predicted 10MHz, reality: 3MHz
  - Store instrs and branch mispredict recovery caused the arith unit to stall. Stalls were more frequent than predicted.
- Legacy
  - IBM 360/91 (codename Project X): 1967
    - very successful and influential (Tomosulo's algorithm)
    - more on this later
  - IBM ACS-1 (codename Project Y): 1963
    - great technology and engineering, ruined by management and market forces
    - more on this later

26

Feranti Pegasus (1956)

- First computer with general-purpose registers
  - 8 registers (R0—R7)
  - R0 was hardwired to 0

27

Burroughs B5000 (1963)

- (First?) stack-based architecture
- First U.S. computer to support virtual memory
- Goal was to support high-level languages (e.g. Algol)

28

CDC 6600 (1964)

- First supercomputer
- First general-purpose load-store computer
- Deeply pipelined
- Supported out-of-order execution with scoreboard

29

IBM 360 (1964)

- Origin of term “Computer Architecture”
  - “Architecture” = instruction set
  - Terminology is abused today
- 360 was family of computers that shared same instruction-set architecture

30

IBM 360/91 (196?—1967)

- Tomasulo’s algorithm for out-of-order execution

31

IBM ACS-1 (1963-1970)

- ACS = “Advanced Computer System”
- AKA “Project Y” from IBM Stretch legacy
- First superscalar processor
  - Amdahl postulated that max instrs per cycle = 1
  - Cocke decided to disprove Amdahl’s postulate
  - Out-of-order issue
    - based on data-dependencies (contender stack)
    - designed by Lynn Conway of Conway-Mead fame

32

IBM ACS-1

Management and markets

- ACS was architecturally incompatible with IBM's family of 360 computers
- 1968: Amdahl proposed major redesign to be ACS-360
  - 5 years into project, 2 years from delivery
- Project fell into disarray, half of the design team quit
- Other problems
  - Circuit problems
  - Company politics
  - Performance of competing IBM products
  - Economic slowdown
- ACS cancelled in 1969
- Superscalar processors stagnant until mid 1980s

33

Mainframes

- IBM 7094
- CDC 6600
- Digital PDP-6
  - PDP = Programmed Data Processor

34

Minicomputers

- A computer that was sufficiently low cost, that a single person could have dedicated access to it, rather than sharing the computer and scheduling access to it.
- Examples:
  - IBM 1620
  - Royal McBee RPG 4000
  - Digital PDP-5, 8
  - Digital PDP-11
    - An entire family of computers
    - Same assembly language
    - Wide span of prices and performance
    - Large PDP-11s evolved into VAX (Virtual Address Extension)

35

Apollo DN10000 (????)

- A step toward explicit parallelism in ISA
- Dual issue (Int/Fp)
- Each integer instruction has "companion bit" saying whether next instruction is floating-point instruction that can be issued in parallel

36

CISC

- With magnetic core memory, performance bottleneck was fetching instruction from memory
- Each instruction was implemented as a sequence of microinstructions (microcode)
- Microcode stored in ROM
- Access to ROM much faster than access to main memory
- Faster to execute a single complex instruction than a sequence of simple instructions
- Goal was to keep compilation simple by reducing the semantic gap between programming languages and assembly language
- Number of opcodes and addressing modes was advertising feature
- Simple, common instructions were short

Microprocessors

- Microprocessor: single chip processor
- Intel 4004
- 1970s
  - Developed by semiconductor companies (e.g. Intel), not computer companies (e.g. IBM, Digital)
  - Architectures of single-chip microprocessors mimicked architecture of multi-chip minicomputer processors
  - Microcode ROM used huge area on microprocessors
  - Little remaining area for performance optimizations (e.g. pipelining)

Berkeley RISC I

- 1980 Patterson and Ditzel: “The case for the reduced instruction set computer”
  - Architecture for multi-chip processor not necessarily optimal for single-chip microprocessor
  - Simple, fixed length instructions
  - Load/store architecture
  - Unified registers
  - Simplifies pipelining, other optimizations
  - Minimal impact on compilers
- Berkeley RISC I: graduate course project
  - Demonstrated that minimal design effort and expense could achieve performance comparable to state-of-the art CISC microprocessors

Berkeley RISC I

Architectural features

- Register windows
  - Programmer has access to window of 32 registers
  - Processor implements many more
  - When do subroutine call, register window moves
  - Goal is to speed up parameter passing for subroutines
  - Adopted by SUN

RISC

- 1980 memory speed vs processor speed
  - memory access rate 3MHz for random access
  - CISC microprocessor max 2MHz mem access rate
    - In 1970s memory speed increased faster than microprocessor speed
    - CISC microprocessors fabricated in NMOS to keep area reasonable, but sacrificed speed.
  - Memory access was not the bottleneck!

Furber, 2000, pp26

ARM

- Acorn RISC Machine
- Acorn Computers Limited was a UK computer company, evolved into Advanced Risc Machines Limited
- The Acorn Risc Machine was the first commercial RISC microprocessor (1983)
  - Later renamed to ***Advanced*** RISC Machine
- Focus is on small, inexpensive, low-power RISC processors
- ARM sells designs, not implementations

Stanford MIPS (1981)

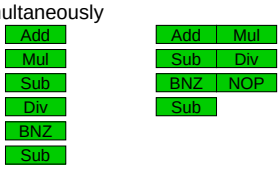
- Led by David Patterson
- “MIPS” = Microprocessor without interlocking pipeline stages
- Goal:
  - Keep the hardware simple
  - Allows increased clock speed
  - Requires compiler to do more work

Pre-RISC RISC

- IBM 801 (1975)
  - Led by John Cocke
  - Experimental computer to explore RISC ideas
- Digital PDP-8
- Cray 1

Computer Architecture in the 1980s

- VLIW: Very long instruction word
  - Attempt to get IPC > 1
  - Use compiler to statically schedule instructions into groups of instructions that can be executed simultaneously



- Rebirth of superscalar
  - IBM Cheetah (early 1980s)
  - IBM America (1985--86)
  - IBM RS/6000 (1989--94)

IBM Cheetah and America

Cheetah

- Started in early 1980s
- Tilak Agerwala coined term “superscalar” in 1983
- 2-wide superscalar
- Research and development experiment, not intended to be a product

America (1985—1986)

- Continuation of Cheetah
- Moved from IBM TJ Watson in New York to Austin Texas
- 4-wide superscalar

IBM RS/6000 (1989 – 1994)

- Commercial introduction 1989
- Evolved into IBM Power family
  - High-end processor for compute servers
  - Continues today as Power5
- IBM Power has offshoot as PowerPC
  - Combined effort with Motorola 1992—2000
  - PowerPC development now down by IBM Austin

Multiflow

- Dataflow architecture: attempt to make parallel nature of hardware visible in instruction set
- Much academic work (Arvind at MIT)
- Failed commercial company: Multiflow
- After demise of Multiflow, some key architects joined Intel to develop P6 processor (aka Pentium Pro, Pentium II)

Explicit Parallelism

- VLIW exposes many implementation details in ISA
- Positive:
- Negative:
- Explicit parallelism lets compiler tell hardware that groups of instructions are data-independent in an implementation-independent manner

UOP / ROP

- UOP (Intel) = micro operation
- ROP (AMD) = RISC operation
- Intel and AMD processors translate each IA-32 instruction into one or more simple, RISC like instructions
- The back end of a CISC processor is almost indistinguishable from a RISC processor
- Previous incarnations:
  - Yale Patt HPS 1986
  - David Ditzel AT&T Crips 1987
- Post incarnations:
  - IBM Power 4

Reinvention of Technology

Simultaneous Multi-Threading

- A single processor runs multiple threads simultaneously, appearing as multiple processors to software
- Previous incarnation
  - IBM ACS-360 simulator
- First reincarnations
  - Intel Pentium 4
  - IBM Power5
- Current processors
  - Sun Niagara
  - Intel Core-2
  - IBM Power 6
  - IBM Cell
  - IBM z6

Multi-Processing

- 1990 – 1996: Braniac vs Speed demon
- 1996 – 2004: Clock speed = performance
- 2004 -- ????: Multicore
  - multiple processors on a single core
  - key difference from past multiprocessing is fast communication between processors
- Design options and challenges
  - how many cores?
  - homogeneous or heterogenous?
  - software?
  - view.eecs.berkeley.edu

56

Schedule

01	Intro and Overview	In-order execution
02	Performance; Architecture	
03	Pipelining review	
04	Superscalar execution	O-o-O execution
05	Out-of-order execution: ctrl and data	
06	Out-of-order execution: reordering	
07	PowerPC 620 vs Intel P6	Current practice
08	Processor Survey	
09	Processor Trends	
10	Data reuse and speculation	Recent advances
11	Multithreading concepts	
12	Multithreading techniques	
13	Review	

60

Administrivia

59