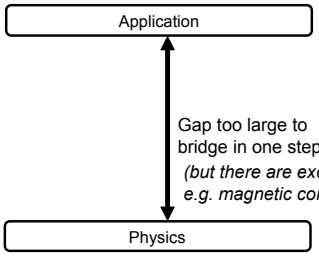
CSE 490/590 Computer Architecture

Introduction

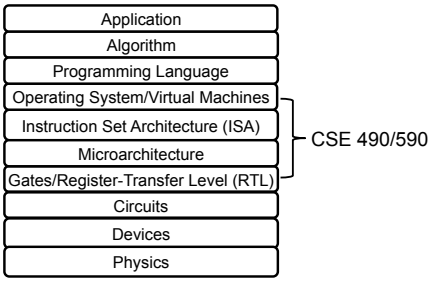
Steve Ko
Computer Sciences and Engineering
University at Buffalo

What is Computer Architecture?



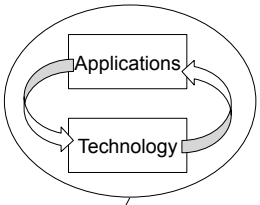
In its broadest definition, computer architecture is the *design of the abstraction layers* that allow us to implement information processing applications efficiently using available manufacturing technologies.

Abstraction Layers in Modern Systems

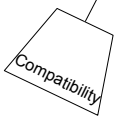


Architecture continually changing

Applications suggest how to improve technology, provide revenue to fund development

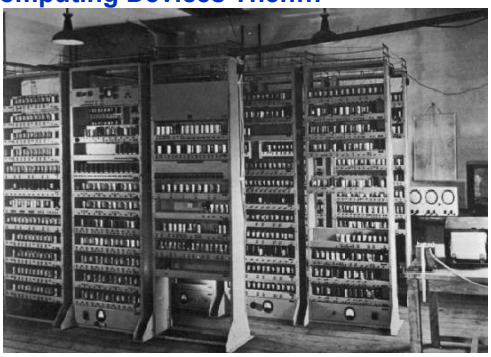


Improved technologies make new applications possible



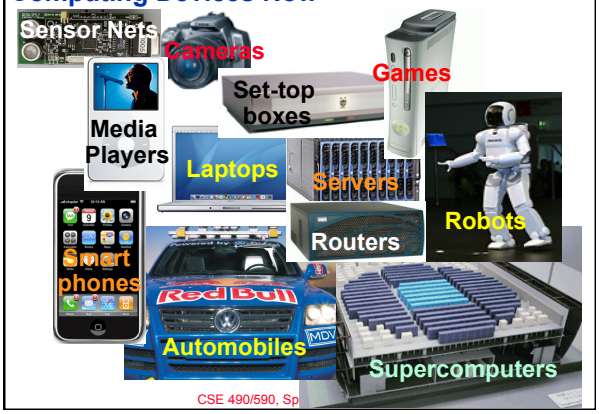
Cost of software development makes compatibility a major force in market

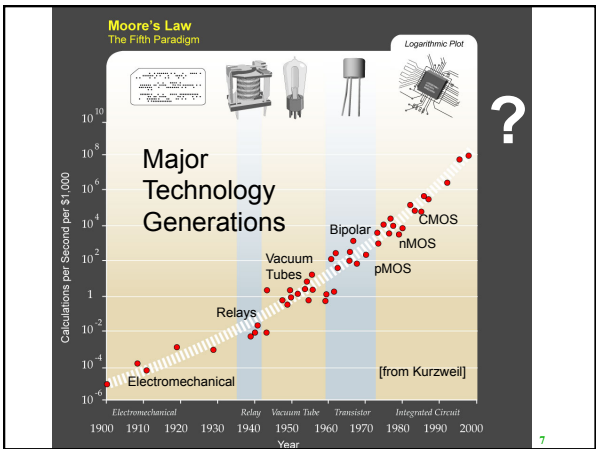
Computing Devices Then…



EDSAC, University of Cambridge, UK, 1949

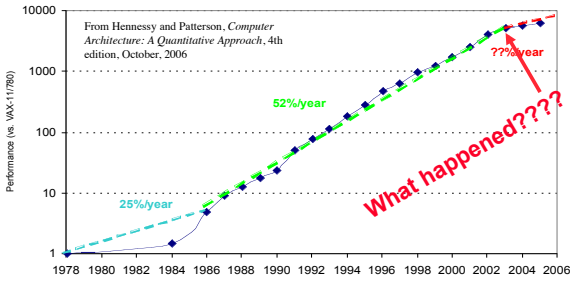
Computing Devices Now





7

Uniprocessor Performance



- VAX : 25%/year 1978 to 1986
- RISC + x86: 52%/year 1986 to 2002
- RISC + x86: ??%/year 2002 to present

The End of the Uniprocessor Era

Single biggest change in the history of computing systems

Crossroads: Conventional Wisdom in Comp. Arch

- Old Conventional Wisdom: Power is free, Transistors expensive
- New Conventional Wisdom: "Power wall" Power expensive, Xtors free (Can put more on chip than can afford to turn on)
- Old CW: Sufficiently increasing Instruction Level Parallelism via compilers, innovation (Out-of-order, speculation, VLIW, …)
- New CW: "ILP wall" law of diminishing returns on more HW for ILP
- Old CW: Multiplies are slow, Memory access is fast
- New CW: "Memory wall" Memory slow, multiplies fast (200 clock cycles to DRAM memory, 4 clocks for multiply)
- Old CW: Uniprocessor performance 2X / 1.5 yrs
- New CW: Power Wall + ILP Wall + Memory Wall = Brick Wall
  - Uniprocessor performance now 2X / 5(?) yrs

⇒ Sea change in chip design: multiple "cores" (2X processors per chip / ~ 2 years)
» More simpler processors are more power efficient

Déjà vu all over again?

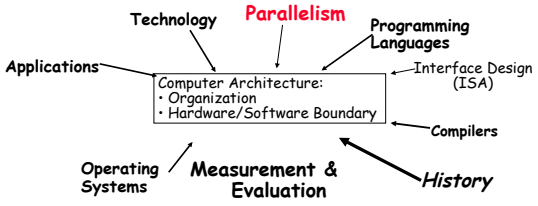
- Multiprocessors imminent in 1970s, '80s, '90s, …
- "… today's processors … are nearing an impasse as technologies approach the speed of light.."
  David Mitchell, *The Transputer: The Time Is Now* (1989)
- Transputer was premature
  ⇒ Custom multiprocessors strove to lead uniprocessors
  ⇒ Procrastination rewarded: 2X seq. perf. / 1.5 years
- "We are dedicating all of our future product development to multicore designs. … This is a sea change in computing"
  Paul Otellini, President, Intel (2004)
- Difference is all microprocessor companies switch to multiprocessors (AMD, Intel, IBM, Sun; all new Apples 2 CPUs)
  ⇒ Procrastination penalized: 2X sequential perf. / 5 yrs
  ⇒ Biggest programming challenge: 1 to 2 CPUs

Problems with Sea Change

- Algorithms, Programming Languages, Compilers, Operating Systems, Architectures, Libraries, … not ready to supply Thread Level Parallelism or Data Level Parallelism for 1000 CPUs / chip,
- Architectures not ready for 1000 CPUs / chip
  - Unlike Instruction Level Parallelism, cannot be solved by just by computer architects and compiler writers alone, but also cannot be solved *without* participation of computer architects
- This 4th Edition of textbook Computer Architecture: A Quantitative Approach explores shift from Instruction Level Parallelism to Thread Level Parallelism / Data Level Parallelism

CSE 490/590 Course Focus

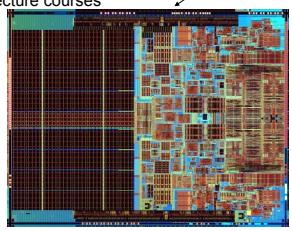
Understanding the design techniques, machine structures, technology factors, evaluation methods that will determine the form of computers in 21st Century

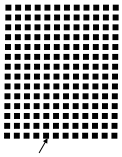


CSE 490/590 Executive Summary

The processor of focus in traditional architecture courses

What you'll understand and experiment with in CSE 490/590





Plus, the technology behind chip-scale multiprocessors (CMPs)

CSE 490/590 Administrivia

Instructor: Steve Ko
Office: 210 Bell Hall, **stevko@buffalo.edu**
Office Hours: Wed. 9am-12PM

TAs: Safwan Weshah, **srwshah@buffalo.edu**
Jangyoung Kim, **jangyoun@buffalo.edu**
Office Hours: TBD

Lectures: MWF, 12pm-12:50pm, 97 Alumni

Recitations: M 9am-9:50am, W 10am-10:50am, F 11am-11:50am

Text: *Computer Architecture: A Quantitative Approach, 4th Edition* (Oct, 2006)

Web page: **http://www.cse.buffalo.edu/~stevko/courses/cse490/spring11**
Lectures available online the morning before class

CSE 490/590 Structure and Syllabus

(Tentative) Five modules

1. Simple machine design (ISAs, microprogramming, unpipelined machines, Iron Law, simple pipelines)
2. Memory hierarchy (DRAM, caches, optimizations) plus virtual memory systems, exceptions, interrupts
3. Complex pipelining (score-boarding, out-of-order issue)
4. Explicitly parallel processors (vector machines, VLIW machines, multithreaded machines)
5. Multiprocessor architectures (cache coherence, memory models, synchronization)

CSE 490/590 Course Components

- 2 Quizzes (20%)
- 2 Exams
  - In-class, closed-book, no calculators or computers.
  - Based on lectures and problem sets
  - Midterm 20%
  - Final 25%
- 35% Projects
  - One project to get you familiarized with the BASYS2 board (5%)
  - Another more substantial project you can choose from a list (30%)
  - The list will be up before the project 1 deadline.

Recitations & HW Assignments

- It is very, very important to attend the recitations.
- Why?
  - For the first 5 weeks, we will cover Verilog and how to use BASYS2 board.
  - This is different from previous offerings.
- To counter the load, I will slow down in the beginning.
- Projects are a big part of this course.
- There will be homework assignments, but we will not grade them.
  - The main purpose is to help you understand the materials.

Late Submission & Regrading

- Late submission
  - Submissions are always due in the beginning of the class.
  - Late submissions will result in 20% penalty a day.
  - After 5 days, it'll be 0%.
- Regrading
  - Regrade requests are due no later than 1 week.
  - Regrade requests must be clearly written and attached to the assignment.
  - When submitted, *everything will be regraded*, not just the one you have a question on. This may result in a lower grade.
  - Work done in pencil will not be considered.

Purchase BASYS2 Board!

- You need to purchase it as soon as possible!
- http://www.digilentinc.com/Products/Detail.cfm?Prod=BASYS2
- Get the student version ($49)

And in conclusion …

- Computer Architecture >> ISAs and RTL
- CSE 490/590 is about interaction of hardware and software, and design of appropriate abstraction layers
- Computer architecture is shaped by technology and applications
  - History provides lessons for the future
- Computer Science at the crossroads from sequential to parallel computing
  - Salvation requires innovation in many fields, including computer architecture
- Read Chapter 1 for next time!

Acknowledgements

- These slides heavily contain material developed and copyright by
  - Krste Asanovic (MIT/UCB)
  - David Patterson (UCB)
- And also by:
  - Arvind (MIT)
  - Joel Emer (Intel/MIT)
  - James Hoe (CMU)
  - John Kubiatowicz (UCB)
- MIT material derived from course 6.823
- UCB material derived from course CS252