

ESE534: Computer Organization

Day 1: January 11, 2012
Introduction and Overview



Today

- Matter Computes
- Architecture Matters
- This Course (short)
- Unique Nature of This Course
- Change
- More on this course

Power of Computation

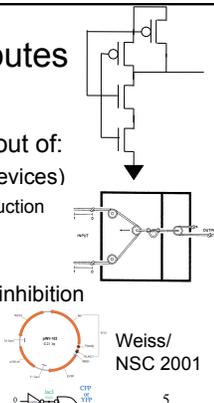
- Which set or gates is more powerful?
 - Set 1: AND2, AND3, AND4
 - Set 2: AND2, OR2
 - Set 3: NAND2
 - Set 4: AND2, XOR2
- (assume have unlimited number of gates in each set)

Review (assert?): Two Universality Facts

- NAND gate Universality [Day 2, ESE170/CIS240]
 - We can implement any computation by interconnecting a sufficiently large network of NAND gates
- Turing Machine is Universal [CIS262]
 - We can implement any *computable* function with a TM
 - We can build a single TM which can be programmed to implement any computable function
- Day 2 reading (on Blackboard) SciAm-level review

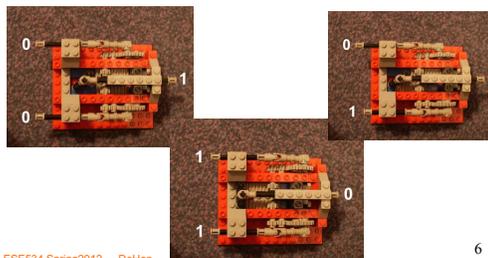
Matter Computes

- We can build NAND gates out of:
 - transistors (semiconductor devices)
 - physical laws of electron conduction
 - mechanical switches
 - basic physical mechanics
 - protein binding / promotion / inhibition
 - Basic biochemical reactions
 - ...many other things



LEGO™ Logic Gates

- <http://goldfish.ikaruga.co.uk/logic.html>



Starting Point



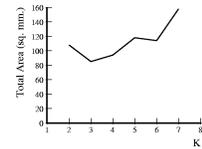
- Given sufficient raw materials:
 - can implement any computable function
- Our goal in computer architecture
 - is **not** to figure out how to compute new things
 - rather, it is an **engineering** problem

Penn ESE534 Spring2012 -- DeHon

7

Engineering Problem

- Implement a computation:
 - with least resources (in fixed resources)
 - with least cost
 - in least time (in fixed time)
 - with least energy
 - With fixed energy budget
- Optimization problem
 - how do we do it best?



Penn ESE534 Spring2012 -- DeHon

8

Quote

- “An Engineer can do for a dime what everyone else can do for a dollar.”

Penn ESE534 Spring2012 -- DeHon

9

How much difference?

- Experience running things on multiple architectures?
 - E.g. GPU, FPGA, Processor....
 - Preferably at same technology node.
 - Same Silicon die area

Penn ESE534 Spring2012 -- DeHon

10

Architecture Matters?

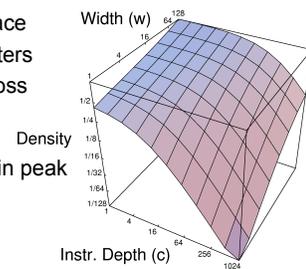
- How much difference is there between architectures?
- How badly can I be wrong in implementing/picking the wrong architecture?
- How efficient is the IA-32, IA-64, GPGPU?
 - Is there much room to do better?
- Is architecture done?
 - A solved problem?

Penn ESE534 Spring2012 -- DeHon

11

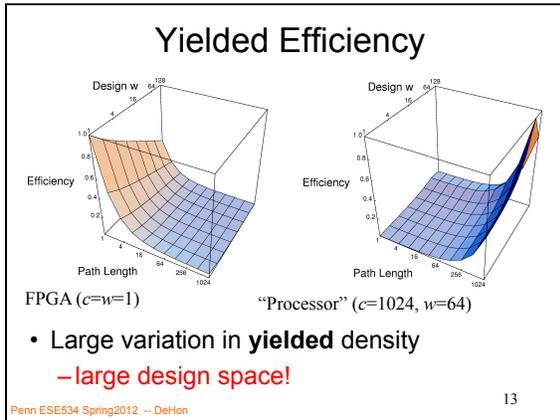
Peak Computational Densities from Model

- Small slice of space
 - only 2 parameters
- 100x density across
- Large difference in peak densities
 - large design space!



Penn ESE534 Spring2012 -- DeHon

12



- ### Architecture **Not** Done
- Many ways, not fully understood
 - design space
 - requirements of computation
 - limits on requirements, density...
 - ...and the costs are **changing**
 - optimal solutions change
 - dominant constraints change
 - creating new challenges and opportunities
- Penn ESE534 Spring2012 -- DeHon 14

- ### Personal Goal?
- Develop systematic design
 - Parameterize design space
 - adapt to costs
 - Understand/capture req. of computing
 - Efficiency metrics
 - (similar to information theory?)
 - ...we'll see a start at these this term
-
- Penn ESE534 Spring2012 -- DeHon 15

- ### Architecture **Not** Done
- Not here to just teach you the forms which are already understood
 - (though, will do that and give you a strong understanding of their strengths and weaknesses)
 - **Goal:** enable you to design and synthesize new and better architectures
- Penn ESE534 Spring2012 -- DeHon 16

- ### This Course (short)
- How to organize computations
 - Requirements
 - Design space
 - Characteristics of computations
 - Building blocks
 - compute, interconnect, retiming, instructions, control
 - Comparisons, limits, tradeoffs
- Penn ESE534 Spring2012 -- DeHon 17

- ### This Course
- Sort out:
 - Custom, RISC, SIMD, Vector, VLIW, Multithreaded, Superscalar, EPIC, MIMD, FPGA, GPGPUs
 - Basis for design and analysis
 - Techniques
 - [more detail at end]
- Penn ESE534 Spring2012 -- DeHon 18

Graduate Class

- Assume you are here to learn
 - Motivated
 - Mature
- Reading
 - Not 1:1 with lecture and assignments
 - Won't be policing you
 - You may need to follow some links beyond "required" reading
- Problems
 - May not be fully, tightly specified

Penn ESE534 Spring2012 -- DeHon

19

Uniqueness of Class

Penn ESE534 Spring2012 -- DeHon

20

Not a Traditional Arch. Class

- Traditional class (240, 370, 501)
 - focus RISC Processor
 - history
 - undergraduate class on μ P internals
 - then graduate class on details
- This class
 - much broader in scope
 - develop design space
 - see RISC processors in context of alternatives

Penn ESE534 Spring2012 -- DeHon

21

Authority/History

- "Science is the belief in the ignorance of experts."
-- Richard Feynman
- Traditional Architecture has been too much about history and authority
- Should be more about engineering evaluation
 - **physical world** is "final authority"
- **Goal:** Teach you to think critically and independently about computer design.



Penn ESE534 Spring2012 -- DeHon

22

Next Few Lectures

- Quick run through logic/arithmetic basics
 - make sure everyone remembers
 - (some see for first time?)
 - get us ready to start with observations about the key components of computing devices
- Trivial/old hat for many
 - But will be some observations couldn't make in ESE170/CIS370
- May be fast if seeing for first time
- Background quiz intended to help me tune

Penn ESE534 Spring2012 -- DeHon

23

Themes

- Design Space
- Parameterization
- Costs
- Change
- Structure in Computations

Penn ESE534 Spring2012 -- DeHon

24

Focus

- Focus on raw computing organization
- **Not** worry about nice abstractions, models
 - 501, 370, 240 provide a few good models
 - Instruction Set Architecture (ISA)
 - Shared Memory
 - Transactional...
 - ...and you should know others
 - Dataflow, streaming, data parallel, ...

Penn ESE534 Spring2012 -- DeHon

25

Change

- A key feature of the computer industry has been **rapid and continual change**.
- We must be prepared to adapt.
- True of this course as well
 -things are still changing...
 - We'll try to figure it out together...

Penn ESE534 Spring2012 -- DeHon

26

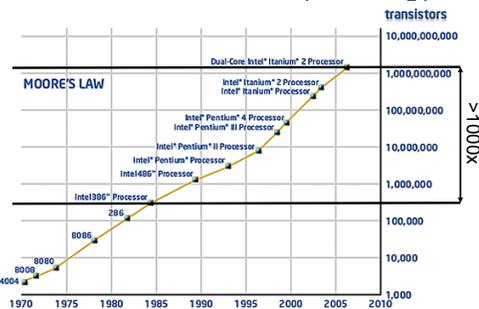
What has changed?

- **[Discuss]**
- Capacity
 - Total
 - Per die
- Size
- Applications
 - Number
 - Size/complexity of each
 - Types/variety
- Use Environment
 - Embedded
 - Mission critical
- Speed
 - Ratio of fast memory to dense memory
 - Wire delay vs. Gate delay
 - Onchip vs. inter-chip
- Joules/op
- Mfg cost
 - Per transistor
 - Per wafer
 - NRE (Non-recurring engineering)
- Reliability
- Limited by
 - Transistors, energy...

Penn ESE534 Spring2012 -- DeHon

27

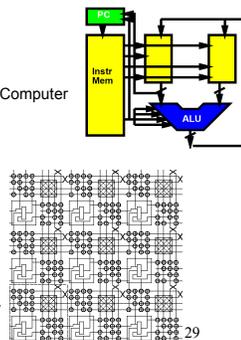
Intel's Moore's Law (Scaling)



Penn ESE534 Spring2012 -- DeHon

1983 (early VLSI)

- Early RISC processors
 - RISC = Reduced Instruction Set Computer
 - RISC-II, 40K transistors
 - MIPS, 24K transistors
 - ~10MHz clock cycle
- Xilinx XC2064
 - 64 4-LUTs
 - LUT = Look-Up Table
 - 4-LUT – program to be any gate of 4 inputs

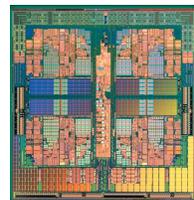


Penn ESE534 Spring2012 -- DeHon

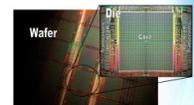
29

- CPUs
 - Billions of transistors
 - 6+ CPU per die
 - Multi-issue, 64b processors
 - GHz clock cycles
 - MByte caches
- FPGAs
 - >500,000 bit processing elements
 - 10s of Mbits of on-chip RAM

AMD Quad Core Opteron



Altera Stratix IV



Penn ESE534 Spring2012 -- DeHon

MOS Transistor *Scaling* (1974 to present)

$$S=0.7$$

[0.5x per 2 nodes]

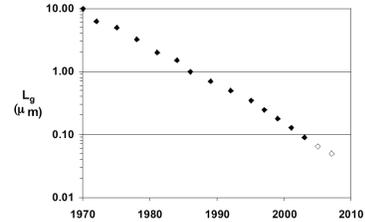


Source: 2001 ITRS - Exec. Summary, ORTC Figure
Penn ESE534 Spring2012 -- DeHon

[from Andrew Kahng]

31

Will This Last Forever?



[Moore, ISSCC2003]

Penn ESE534 Spring2012 -- DeHon

32

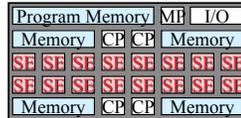
More chip capacity?

Cosmic Cube / CACM 1985

- Should a 2010 single-chip multiprocessor look like a 1983 multiprocessor systems?



- Processor \rightarrow processor latency?
- Inter-processor bandwidth costs?
- Cost of customization?



Calisto™ BCM1500

33

Penn ESE534 Spring2012 -- DeHon

Nichols/Microprocessor Forum 2001

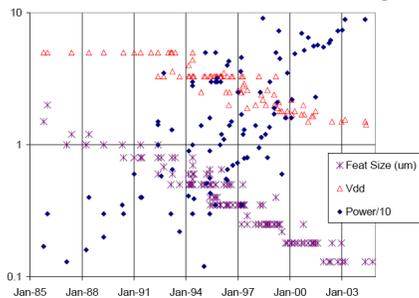
Memory Levels

- Why do we have 5+ levels of memory today?
 - Apple II, IBM PC had 2
 - MIPS-X had 3

Penn ESE534 Spring2012 -- DeHon

34

Historical Power Scaling



Penn ESE534 Spring2012 -- DeHon

[Horowitz et al. / IEDM 2005]

35

Interesting Times

- Challenges to continue scaling
 - Power density
 - Reliability
- What does the end-of-scaling mean to architecture?

Penn ESE534 Spring2012 -- DeHon

36

Class Components

Class Components

- Lecture (incl. preclass exercise)
 - Slides on web before class
 - (you can print if want a follow-along copy)
- Reading [~1 required paper/lecture]
 - No text (mostly online: Blackboard, IEEE, ACM)
- 9 assignments
 - (roughly 1 per week)
- Final design/analysis exercise
 - (~4 weeks)
- [Note syllabus, course admin online](#)

Preclass Exercise

- Like Background Quiz but more focused
- Motivate the topic of the day
 - Introduce a problem
 - Introduce a design space, tradeoff, transform
- Work for ~5 minutes before start lecturing

Feedback

- Will have anonymous feedback sheets for each lecture
 - Clarity?
 - Speed?
 - Vocabulary?
 - General comments

Fountainhead Quote

Howard Roark's Critique of the Parthenon
-- Ayn Rand

Fountainhead Parthenon Quote

"Look," said Roark. "The famous flutings on the famous columns---what are they there for? To hide the joints in wood---when columns were made of wood, only these aren't, they're marble. The triglyphs, what are they? Wood. Wooden beams, the way they had to be laid when people began to build wooden shacks. Your Greeks took marble and they made copies of their wooden structures out of it, because others had done it that way. Then your masters of the Renaissance came along and made copies in plaster of copies in marble of copies in wood. Now here we are making copies in steel and concrete of copies in plaster of copies in marble of copies in wood. Why?"



Penn ESE534 Spring2012 -- DeHon

43

Computer Architecture Parallel

- Are we making:
 - copies in submicron CMOS
 - of copies in early NMOS
 - of copies in discrete TTL
 - of vacuum tube computers?

Penn ESE534 Spring2012 -- DeHon

44

Admin

- Your action:
 - Find course web page
 - Read it, including the policies
 - Find Syllabus
 - Find assignment 1
 - Find lecture slides
 - » Will try to post before lecture
 - Find reading assignments
 - Find reading for lecture 2 on blackboard

Penn ESE534 Spring2012 -- DeHon

45

Big Ideas

- Matter Computes
- Efficiency of architectures varies widely
- Computation design is an engineering discipline
- Costs change \Rightarrow Best solutions (architectures) change
- Learn to cut through hype
 - analyze, think, critique, synthesize

Penn ESE534 Spring2012 -- DeHon

46