

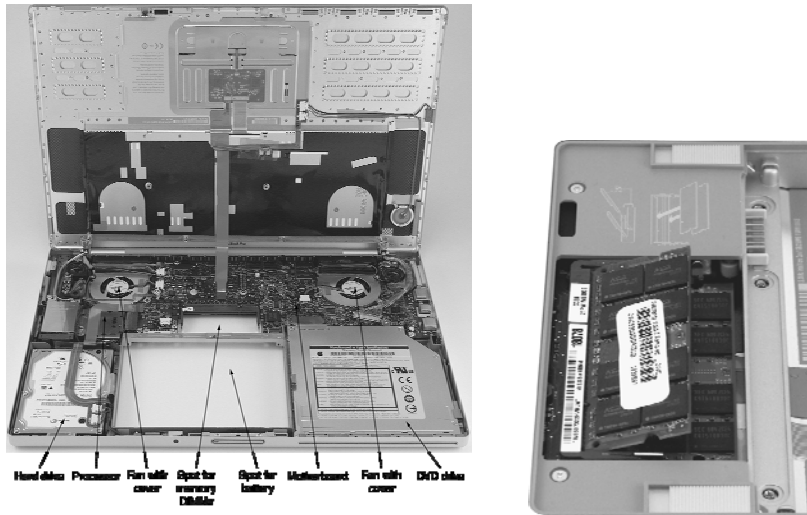
Lec 24: Parallel Processors

Kavita Bala
CS 3410, Fall 2008
Computer Science
Cornell University

Announcements

- PA 3 out
 - Hack ‘n Seek
 - The goal is to have fun with it
 - Recitations today will talk about it
- Pizza party on Tuesday after Thanksgiving
- Final project (distributed ray tracer) out last week
 - Demos: Dec 16 or 17

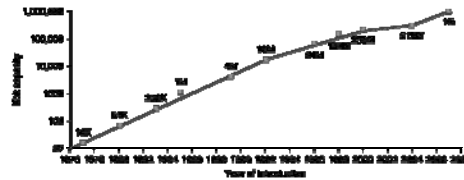
Opening the Box



© Kavita Bala, Computer Science, Cornell University

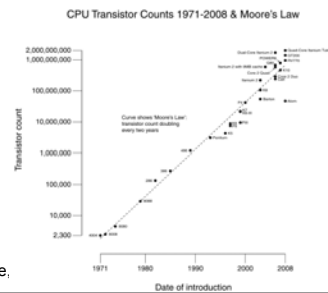
Technology Trends

- DRAM capacity
 - Increased
 - Reduced cost



DRAM capacity

- Moore's Law
 - Speed?
 - Not really

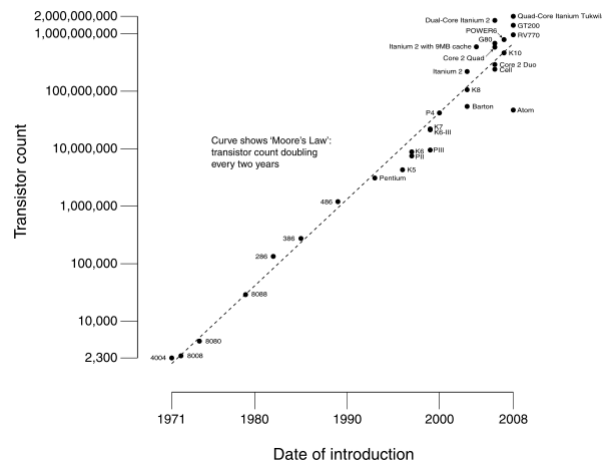


© Kavita Bala, Computer Science,

Moore's Law

- Law about transistor count

CPU Transistor Counts 1971-2008 & Moore's Law



Review: Relative Performance

- Define Performance = 1/Execution Time
- “X is n time faster than Y”

$$\begin{aligned} \text{Performance}_X / \text{Performance}_Y \\ = \text{Execution time}_Y / \text{Execution time}_X = n \end{aligned}$$

- Example: time taken to run a program
 - 10s on A, 15s on B
 - $\text{Execution Time}_B / \text{Execution Time}_A$
 $= 15s / 10s = 1.5$
 - So A is 1.5 times faster than B

Review: CPU Time

$$\begin{aligned}\text{CPU Time} &= \text{CPU Clock Cycles} \times \text{Clock Cycle Time} \\ &= \frac{\text{CPU Clock Cycles}}{\text{Clock Rate}}\end{aligned}$$

- Performance improved by
 - Reducing number of clock cycles
 - Increasing clock rate
 - Hardware designer must often trade off clock rate against cycle count

© Kavita Bala, Computer Science, Cornell University

Review: CPU Time Example

- Computer A: 2GHz clock, 10s CPU time
- Designing Computer B
 - Aim for 6s CPU time
 - Can do faster clock, but causes $1.2 \times$ clock cycles
- How fast must Computer B clock be?

$$\text{Clock Rate}_B = \frac{\text{Clock Cycles}_B}{\text{CPU Time}_B} = \frac{1.2 \times \text{Clock Cycles}_A}{6\text{s}}$$

$$\begin{aligned}\text{Clock Cycles}_A &= \text{CPU Time}_A \times \text{Clock Rate}_A \\ &= 10\text{s} \times 2\text{GHz} = 20 \times 10^9\end{aligned}$$

$$\text{Clock Rate}_B = \frac{1.2 \times 20 \times 10^9}{6\text{s}} = \frac{24 \times 10^9}{6\text{s}} = 4\text{GHz}$$

© Kavita Bala, Computer Science, Cornell University

Review: Instruction Count and CPI

Clock Cycles = Instruction Count \times Cycles per Instruction

CPU Time = Instruction Count \times CPI \times Clock Cycle Time

$$= \frac{\text{Instruction Count} \times \text{CPI}}{\text{Clock Rate}}$$

- Instruction Count for a program
 - Determined by program, ISA and compiler
- Average cycles per instruction
 - Determined by CPU hardware
 - If different instructions have different CPI
 - Average CPI affected by instruction mix

© Kavita Bala, Computer Science, Cornell University

Review: CPI Example

- Computer A: Cycle Time = 250ps, CPI = 2.0
- Computer B: Cycle Time = 500ps, CPI = 1.2
- Same ISA
- Which is faster, and by how much?

$$\begin{aligned} \text{CPU Time}_A &= \text{Instruction Count} \times \text{CPI}_A \times \text{Cycle Time}_A \\ &= I \times 2.0 \times 250\text{ps} = I \times 500\text{ps} \end{aligned}$$

A is faster...

$$\begin{aligned} \text{CPU Time}_B &= \text{Instruction Count} \times \text{CPI}_B \times \text{Cycle Time}_B \\ &= I \times 1.2 \times 500\text{ps} = I \times 600\text{ps} \end{aligned}$$

$$\frac{\text{CPU Time}_B}{\text{CPU Time}_A} = \frac{I \times 600\text{ps}}{I \times 500\text{ps}} = 1.2$$

...by this much

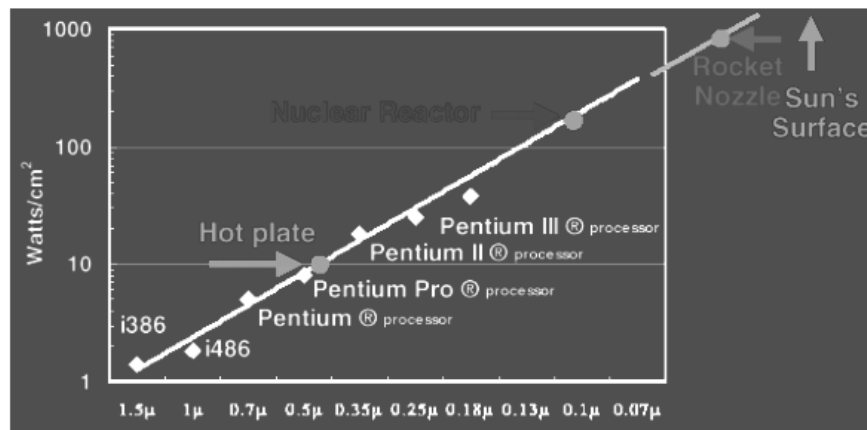
© Kavita Bala, Computer Science, Cornell University

Why Multicore?

- Moore's law
 - A law about transistors
 - Smaller means faster transistors
- Power consumption growing with transistors

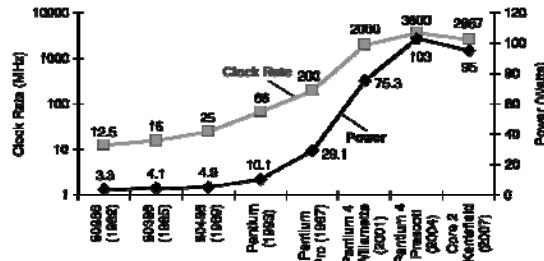
© Kavita Bala, Computer Science, Cornell University

Power Limits Performance



© Kavita Bala, Computer Science, Cornell University

Power Trends



- In CMOS IC technology

$$\text{Power} = \text{Capacitive load} \times \text{Voltage}^2 \times \text{Frequency}$$

×30

5V → 1V

×1000

© Kavita Bala, Computer Science, Cornell University

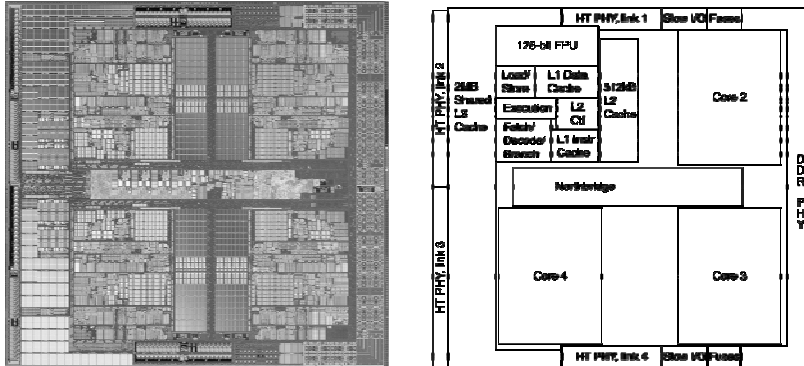
High Power at Idle

- Look back at X4 power benchmark
 - At 100% load: 295W
 - At 50% load: 246W (83%)
 - At 10% load: 180W (61%)
- Google data center
 - Mostly operates at 10% – 50% load
 - At 100% load less than 1% of the time

© Kavita Bala, Computer Science, Cornell University

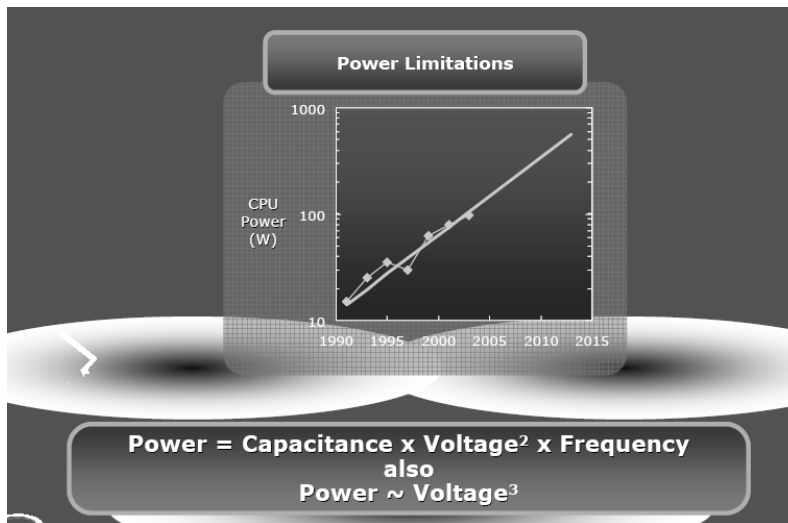
Inside the Processor

- AMD Barcelona: 4 processor cores

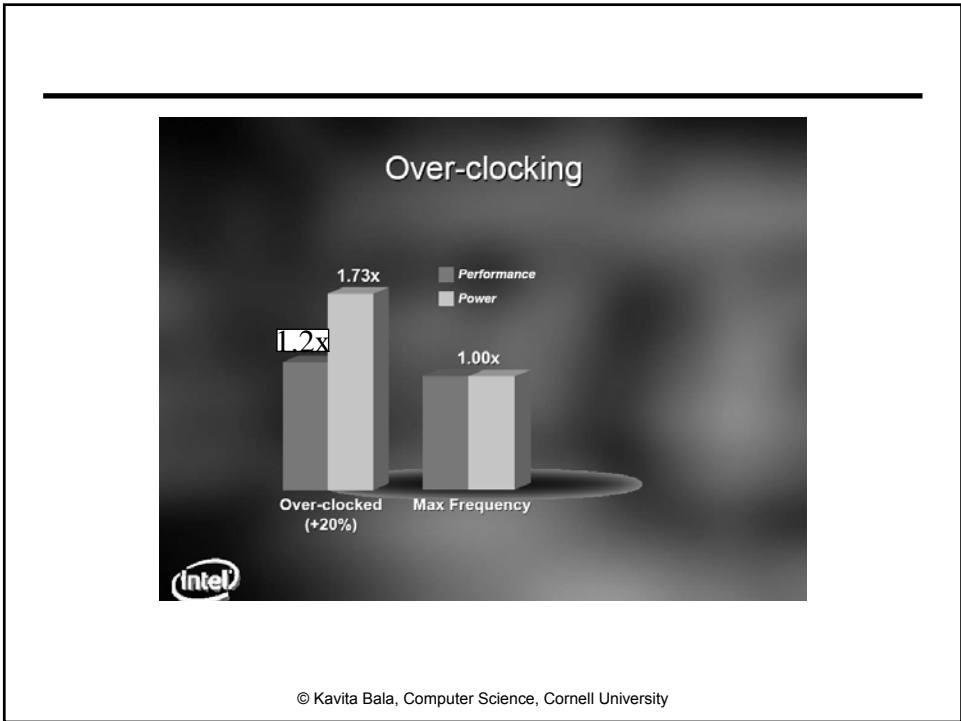
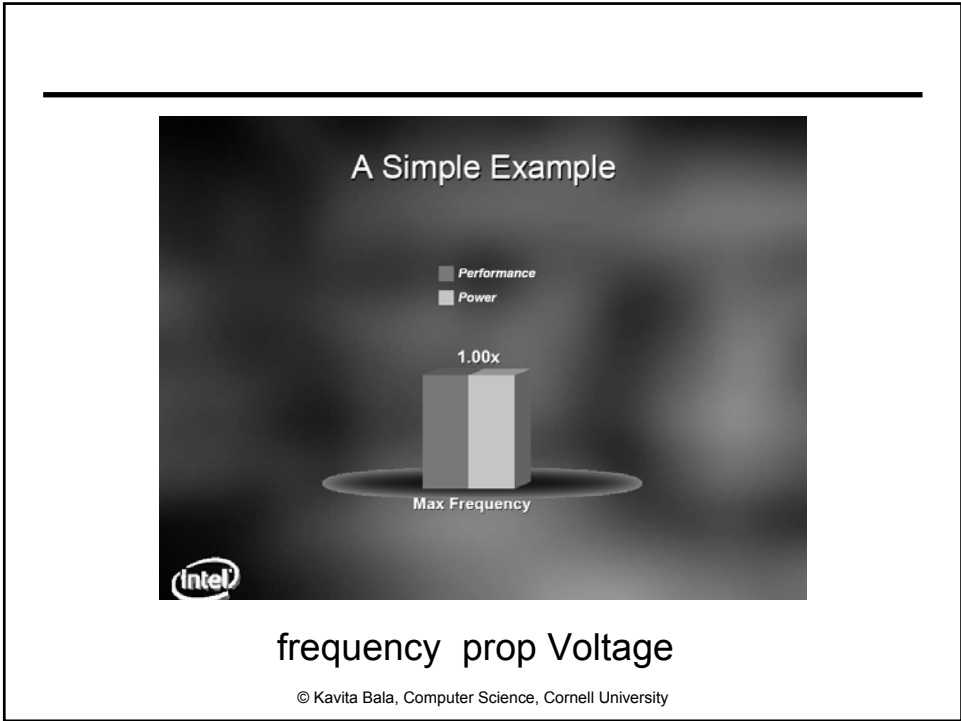


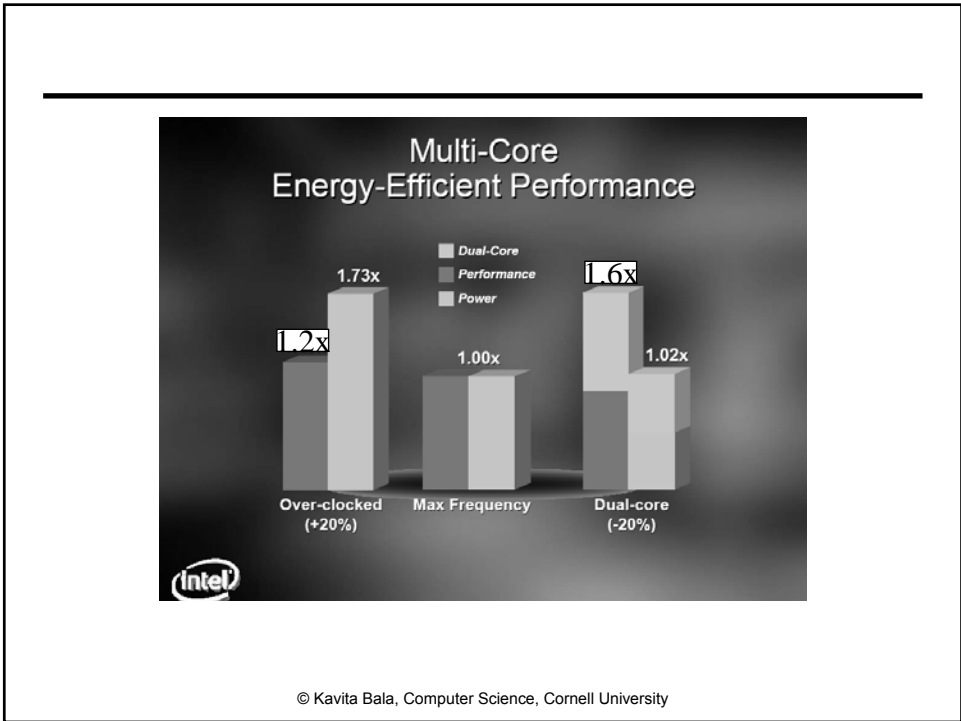
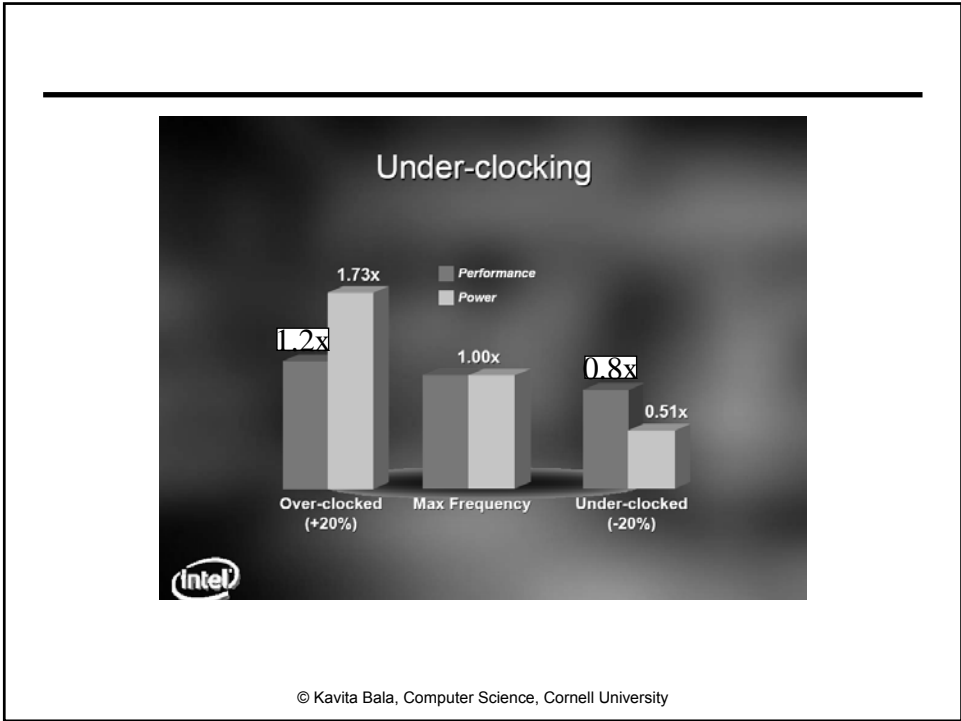
© Kavita Bala, Computer Science, Cornell University

Intel's argument



© Kavita Bala, Computer Science, Cornell University





Multicore: Do you believe?



© Kavita Bala, Computer Science, Cornell University