# Lecture 1: CS/ECE 3810 Introduction

- Today's topics:
  - Why computer organization is important
  - Logistics
  - Modern trends

# Why Computer Organization

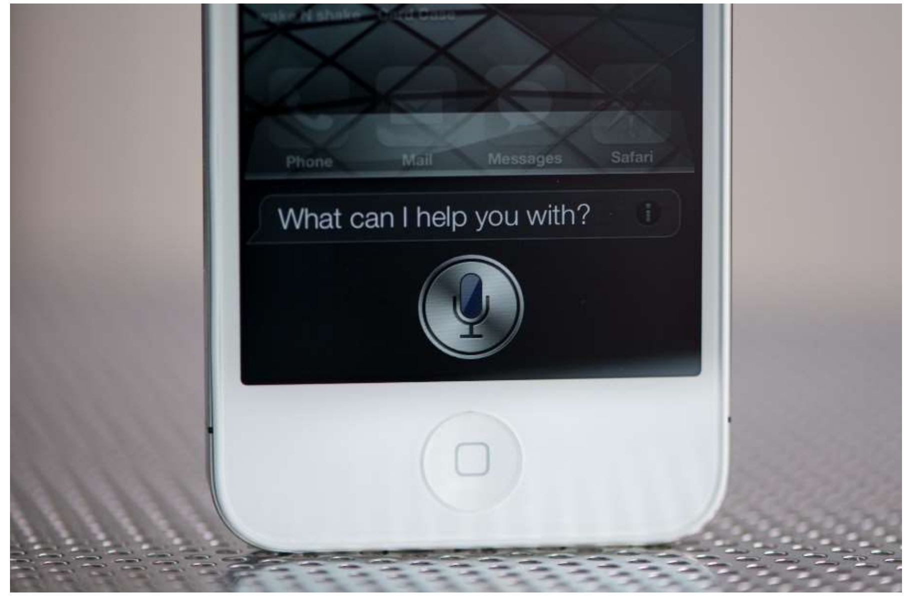




Image credits: uber, extremetech, anandtech

# Why Computer Organization




Image credits: gizmodo

# Why Computer Organization

- Efficiency is important! You learn about big-O notation in Algorithms, you learn about the constants in this class

- The AI revolution owes a lot to fast hardware; a complete CS/CE education includes a deep understanding of hardware

(1) Better algos

(2) Access to datasets (large + labeled)

(3) Access to HW

# Why Computer Organization

- Embarrassing if you are a BS in CS/CE and can’t make sense of the following terms: DRAM, pipelining, cache hierarchies, I/O, virtual memory, …

- Embarrassing if you are a BS in CS/CE and can’t decide which processor to buy: 4.4 GHz Intel Core i9 or 4.7 GHz AMD Ryzen 7 (reason about performance/power)

- Obvious first step for chip designers, compiler/OS writers

- Will knowledge of the hardware help you write better and more secure programs?

# Must a Programmer Care About Hardware?

- Must know how to reason about program performance and energy and security
- Memory management: if we understand how/where data is placed, we can help ensure that relevant data is nearby
- Thread management: if we understand how threads interact, we can write smarter multi-threaded programs

→ Why do we care about multi-threaded programs?

multi-core

# Example

200x speedup for matrix vector multiplication

- Data level parallelism: 3.8x
- Loop unrolling and out-of-order execution: 2.3x
- Cache blocking: 2.5x
- Thread level parallelism: 14x

Further, can use accelerators to get an additional 100x.

# Key Topics

- Moore's Law, power wall
- Use of abstractions
- Assembly language
- Computer arithmetic
- Pipelining
- Using predictions
- Memory hierarchies
- Accelerators
- Reliability and Security

# Logistics

- See class web-page for syllabus/resources
  https://www.cs.utah.edu/~rajeev/cs3810
- TAs and office hours: on the class webpage
- Most communication on Canvas and Piazza; email me directly to set up meetings, or meet me in office hours right before class
- Textbook: Computer Organization – HW/SW Interface, Patterson and Hennessy, $5^{th}$ or $6^{th}$ edition

# Course Organization

- 40% two midterms, 40% final, 20% assignments
- ~10 assignments – you may skip two; automatic 1.5 day extension until Wed/Fri late night; upload on Gradescope; HW4 has extra credit
- Co-operation policy: you may discuss – you may not see someone else's written matter when writing your solution
- Exams are open-notes (1 page)
- Print slides just before class
- Screencast YouTube videos

# Grading Policy

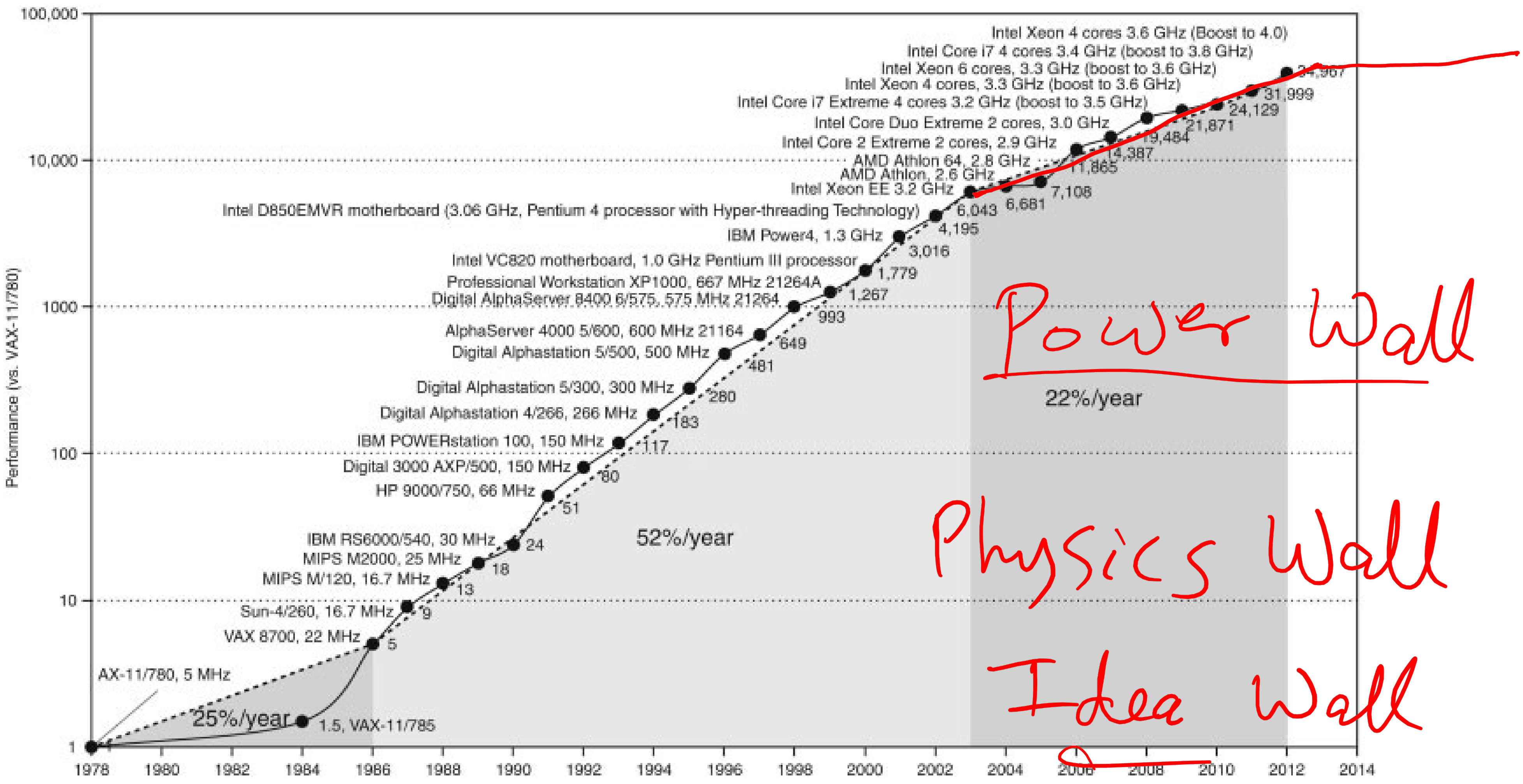
- Grading by rank (32% receive A or A-, 40% receive B+/B/B-)
- About 13% receive D+ or lower — 60 pts/100
- No tolerance for cheating (see class webpage)
- Rank in exams matters more than rank in homeworks

16% A (39 students) 93 87

16% A- 90

# Microprocessor Performance

120W power




Source: H&P Textbook

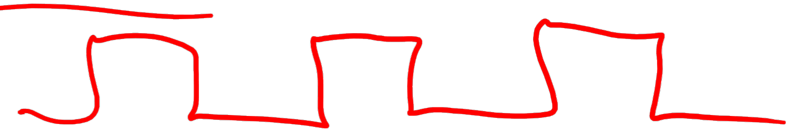
50% improvement every year!!

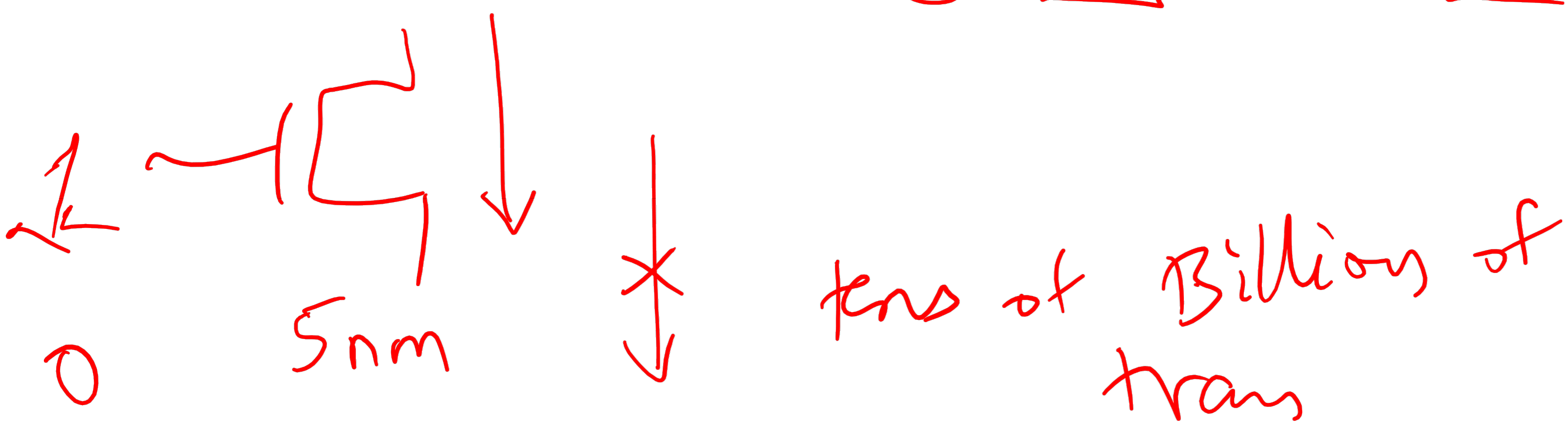
What contributes to this improvement?

Why the lower improvement?

# Transistors and Clocks

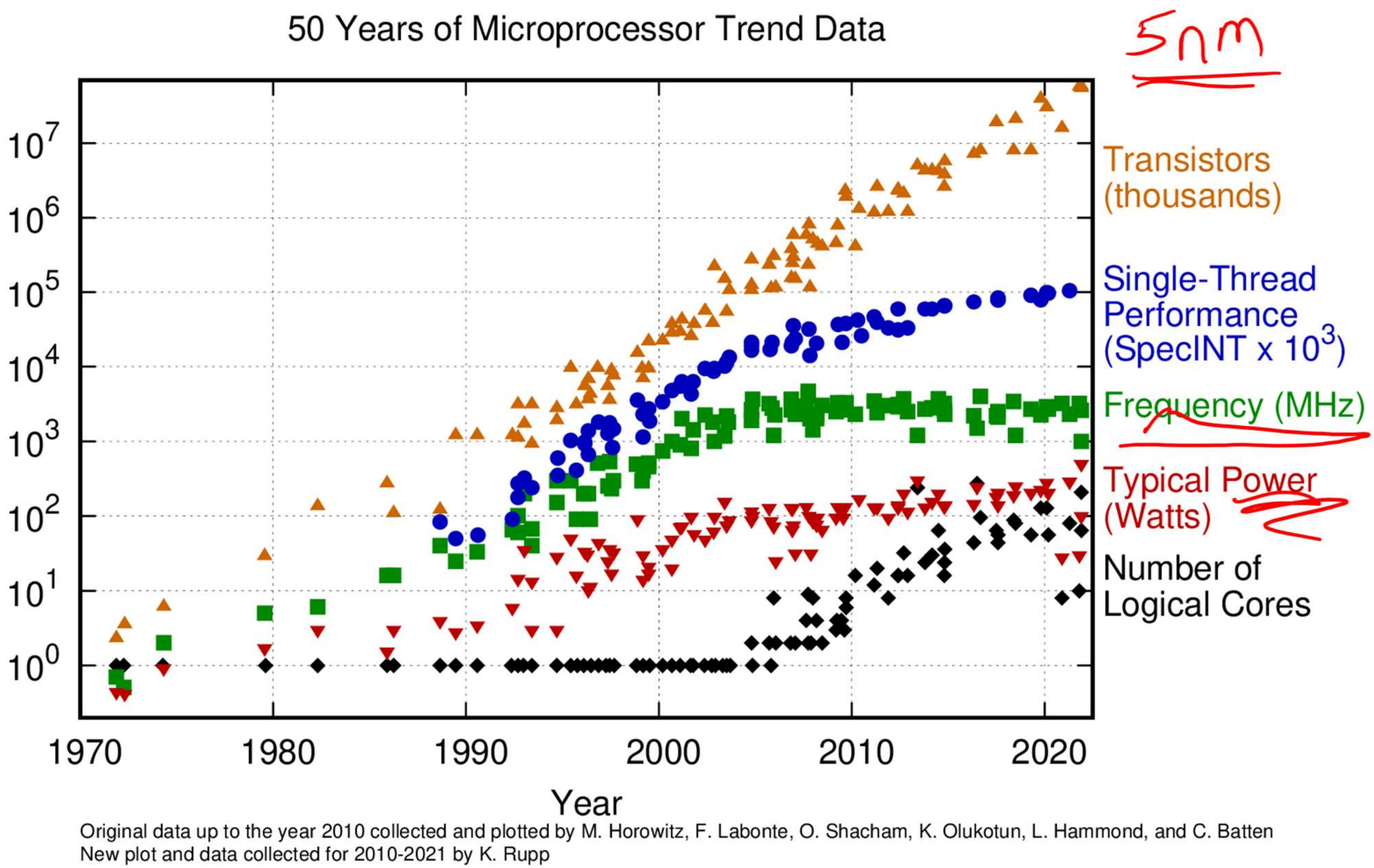
- Transistor: basic device used to create logic gates, circuits
- Smaller transistors → more circuits, faster circuits
- Clock: a voltage signal that goes up and down, serving as a start/stop signal for circuits; 4 GHz clock frequency = 4 billion up/downs in a second





# Microprocessor Performance

chip 2x larger
cost 4x






Source: karlrupp.net

# Microprocessor Performance

Thermal Design point

| Processor branding | Model | Cores (threads) P | Cores (threads) E | Clock rate (GHz) Base P | Base E | Turbo Boost 2.0 P | Turbo Boost 2.0 E | Turbo Boost 3.0 P | TVB P | GPU Model | GPU Max. freq. (GHz) | Smart cache | TDP Base | TDP Turbo | Price (USD)[a] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Core i9 | 14900KS | 8 (16) | 16 (16) | 3.2 | 2.4 | 5.7 | 4.5 | 5.9 | 6.2 | UHD 770 | 1.65 | 36 MB | 150 W | 253 W | $689 |
| Core i9 | 14900K | 8 (16) | 16 (16) | 3.2 | 2.4 | 5.6 | 4.4 | 5.8 | 6.0 | UHD 770 | 1.65 | 36 MB | 125 W | 253 W | $589 |
| Core i9 | 14900KF | 8 (16) | 16 (16) | 3.2 | 2.4 | 5.6 | 4.4 | 5.8 | 6.0 | N/A | | 36 MB | 125 W | 253 W | $564 |
| Core i9 | 14900 | 8 (16) | 16 (16) | 2.0 | 1.5 | 5.4 | 4.3 | 5.6 | 5.8 | UHD 770 | 1.65 | 36 MB | 65 W | 219 W | $549 |
| Core i9 | 14900F | 8 (16) | 16 (16) | 2.0 | 1.5 | 5.4 | 4.3 | 5.6 | 5.8 | N/A | | 36 MB | 65 W | 219 W | $524 |
| Core i9 | 14900T | 8 (16) | 16 (16) | 1.1 | 0.8 | 5.1 | 4.0 | 5.5 | | UHD 770 | 1.65 | 36 MB | 35 W | 106 W | $549 |
| | 14790F | 8 (16) | 8 (8) | 2.1 | 1.5 | 5.3 | 4.2 | 5.4 | | N/A | | 36 MB | 65 W | 219 W | China only |

8 hi-perf

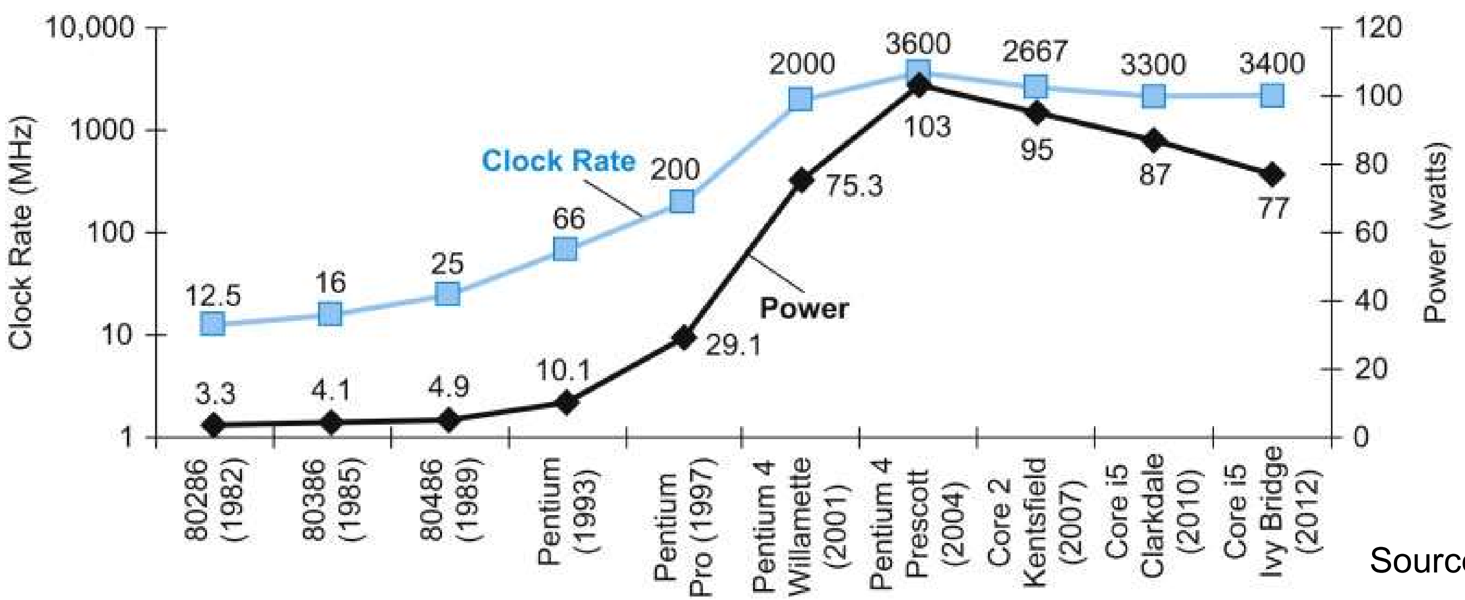
16 energy effic



Source: wikipedia

# Power Consumption Trends

- Dyn power α activity x capacitance x voltage$^2$ x frequency
- Voltage and frequency are somewhat constant now, while capacitance per transistor is decreasing and number of transistors (activity) is increasing
- Leakage power is also rising (function of #trans and voltage)






Source: H&P Textbook

# Important Trends

- Running out of ideas to improve single thread performance
- Power wall makes it harder to add complex features
- Power wall makes it harder to increase frequency

$P \propto f$

- Technology scaling likely to end soon

7nm
5nm

- Additional performance provided by: more cores, occasional spikes in frequency, accelerators

multi-threaded

# Summary

Take-home

- Three roadblocks: power, ideas, technology scaling
- Fixed power budget because of cooling constraints; implies that frequency can't be increased; discourages complex ideas
- End of voltage (Dennard) scaling in early 2010s; the end of Moore's Law also imminent
- Has led to dark silicon and dim silicon (occasional turbo)

Turbo Boost

# Next Class

- Topics: Trends, Performance, MIPS instruction set architecture (Chapter 2)

- Visit the class web-page https://www.cs.utah.edu/~rajeev/cs3810