Computer Architecture

Dr. RAJIB MALL

Professor

Department Of Computer Science & Engineering

IIT Kharagpur.

About Myself

- Name: RAJIB MALL
- B.E. , M.E., Ph.D. from Indian Institute of Science, Bangalore

- Worked with Motorola (India)
  - Senior engineer and later project manager
- Shifted to IIT, Kharagpur in 1994
  - Currently Professor, CSE Dept.

Introduction

- In a first level architecture course --- we first study Von Neumann type of computers:
  - **An instruction is fetched, analyzed, processed, next instruction fetched, analyzed, processed, and so on**
- Ground reality now OfCourse is very different…
- Even the cheapest desktop/laptop/server is:
  - **Multicore**
  - **Superscalar**
  - **Hyperthreaded**
  - **GPU**
  - **Speculative execution**
  - **etc... etc...**

John Luis von Neumann

Von Neumann Stored Program Computers: 1944

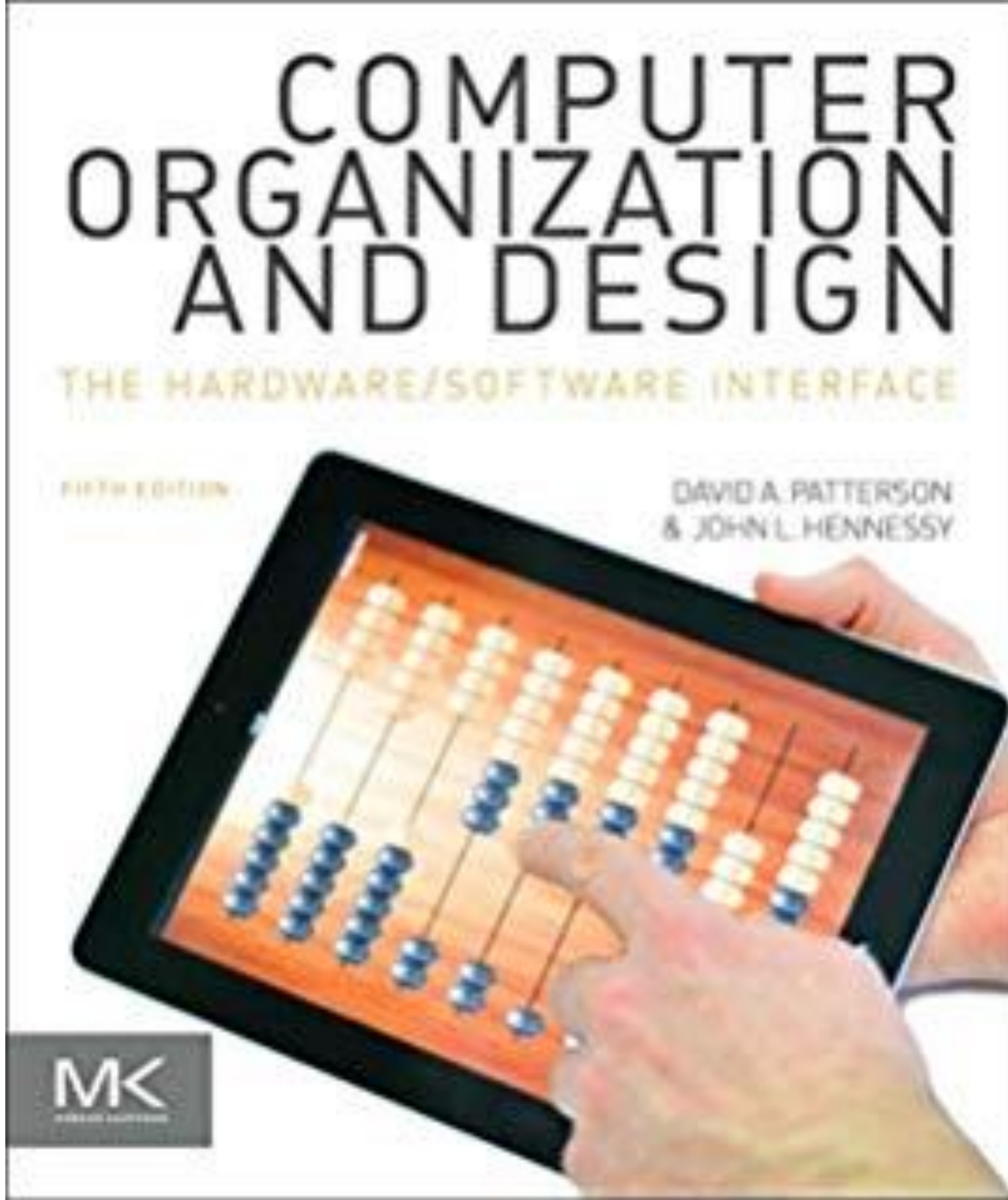
- In older days, "programming" involved flipping switches for every instruction.
  - **Memory stored only data.**

- Around 1944, John von Neumann and others got the idea:
  - **Encode instructions in a format that could be stored in memory just like data.**
  - The processor would interpret and execute the instructions stored in the memory.
  - **Fetch, decode, and execute cycle…**

Grading

- Quizzes every 2 or 3 lectures (Best 75%): 100%
- Class Interactions --- 5%

Text Book


COMPUTER ORGANIZATION AND DESIGN
THE HARDWARE/SOFTWARE INTERFACE
FIFTH EDITION
DAVID A. PATTERSON
& JOHN L. HENNESSY

Decimal term	Abbreviation	Value	Binary term	Abbreviation	Value	% Larger
kilobyte	KB	10^3	kibibyte	KiB	2^{10}	2%
megabyte	MB	10^6	mebibyte	MiB	2^{20}	5%
gigabyte	GB	10^9	gibibyte	GiB	2^{30}	7%
terabyte	TB	10^{12}	tebibyte	TiB	2^{40}	10%
petabyte	PB	10^{15}	pebibyte	PiB	2^{50}	13%
exabyte	EB	10^{18}	exbibyte	EiB	2^{60}	15%
zettabyte	ZB	10^{21}	zebibyte	ZiB	2^{70}	18%
yottabyte	YB	10^{24}	yobibyte	YiB	2^{80}	21%

The Computer Revolution

- Progress in computer technology
  - Captured by Moore's Law
- Has made novel applications feasible
  - Computers in automobiles
  - Cell phones
  - Human genome project
  - World Wide Web
  - 6Search Engines, etc.

Classes of Computers

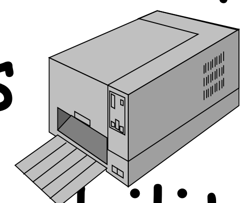
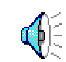
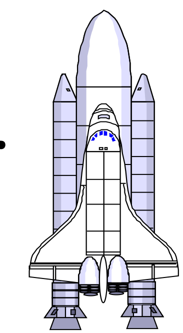
- **Personal computers**
  - General purpose, variety of applications
  - Subject to cost/performance tradeoff

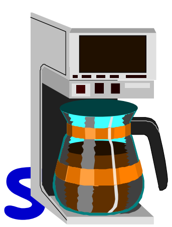
- **Servers**
  - Network based
  - High storage, performance, reliability
  - Range from small servers to building sized

Classes of Computers

- **Supercomputers**
  - High-end scientific and engineering calculations
  - Highest capability but represent a small fraction of the overall computer market

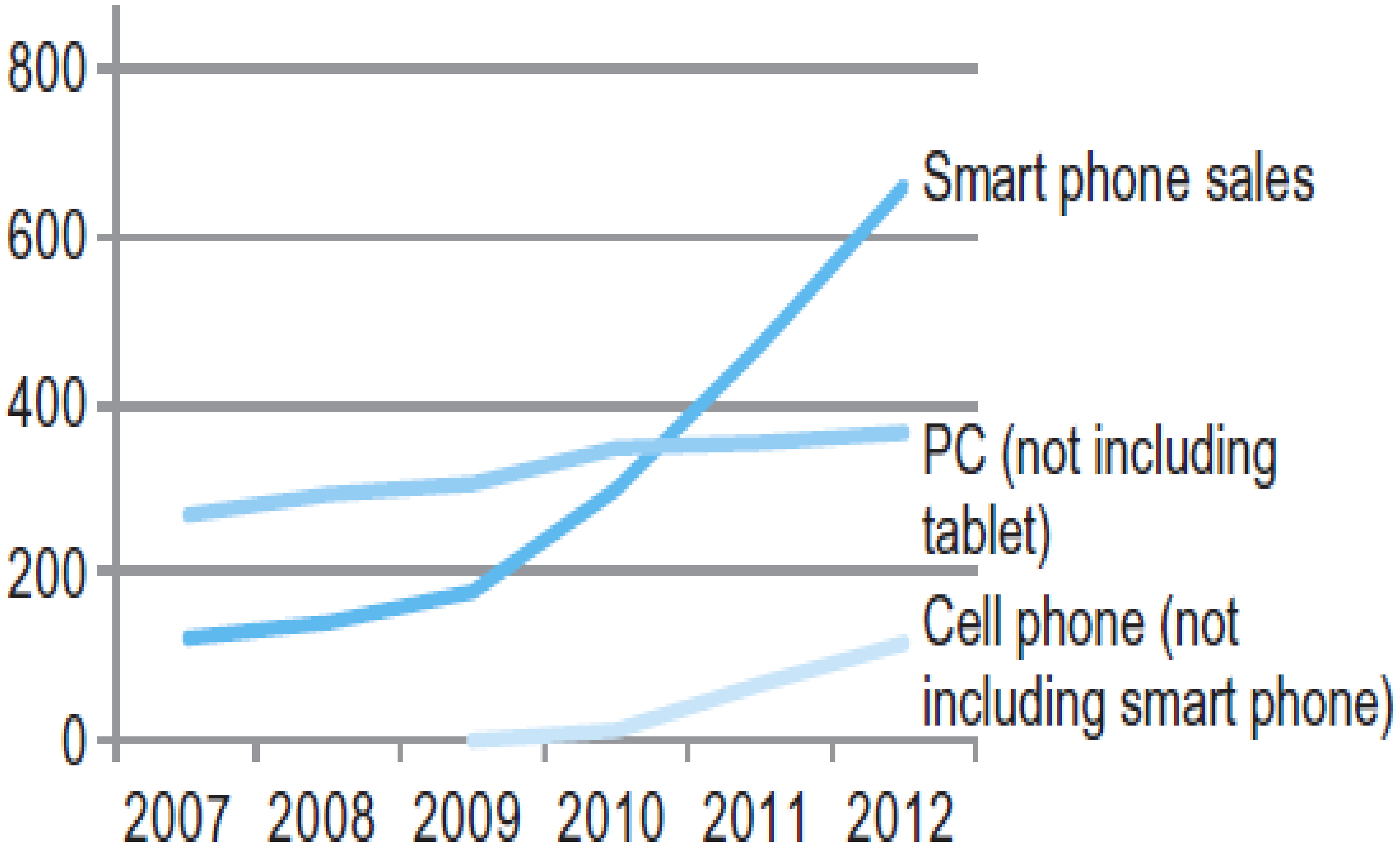


- **Embedded computers**
  - Computer hidden as a part of a system
  - Stringent power/performance/cost constraints

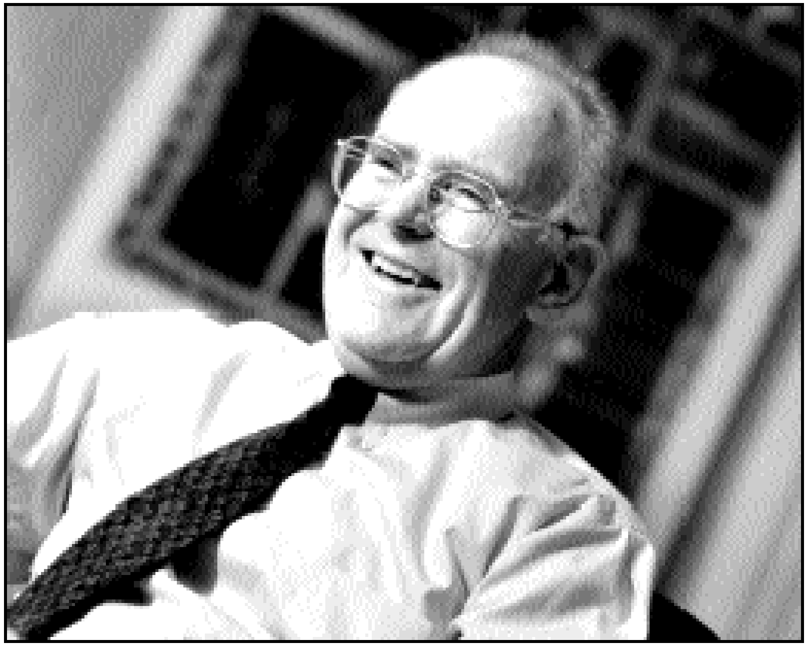
The PostPC Era



Moore's Law

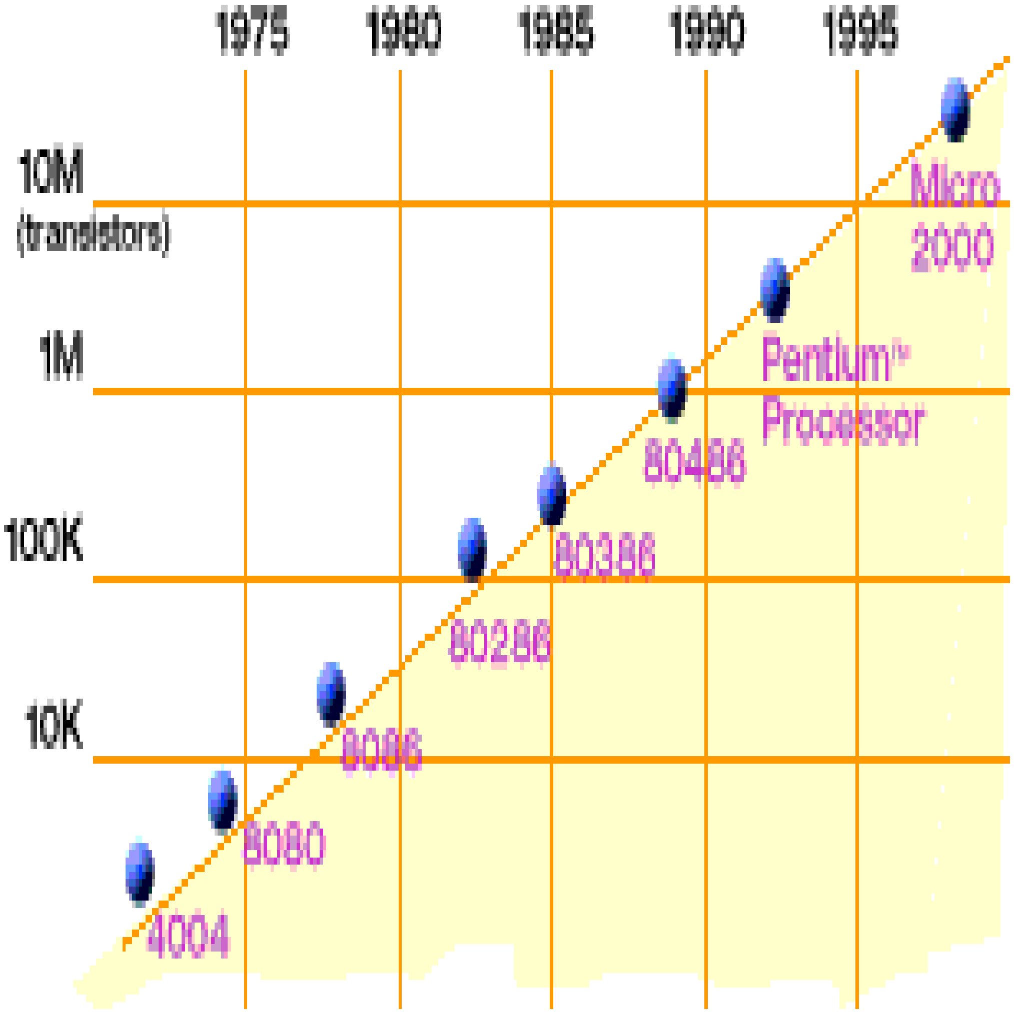
- Computer performance increase over time is captured by Moore's Law (1965):
  - Number of transistors per square inch roughly doubles every eighteen months.
  - Later in 1975 he revised the time to 24 months
- Moore's law is not exactly a law:
  - But, has held good for over 55 years.

Moore's Law



Gordon Moore (co-founder of Intel) predicted in 1965:
"Transistor density of minimum cost semiconductor chips would double roughly every 18 months."

Transistor density is *correlated* to processing speed.



Moore's Law: it has worked for a long time.

Trends Related to Moore's Law

Cont…

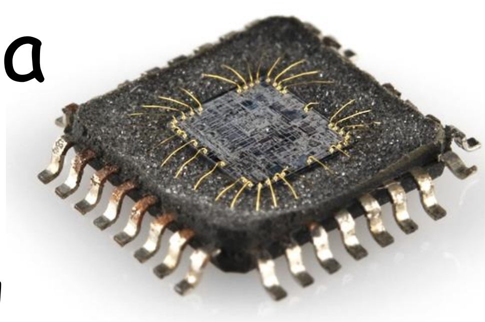
- Processor performance:
  - Twice as fast after every 2 years (roughly).
- Memory capacity:
  - Twice as much after every 18 months (roughly).

Interpreting Moore's Law

- Moore's law is not about just the density of transistors on a chip that can be achieved:
  - **But about the density of transistors at which the cost per transistor is the lowest.**

- As more transistors are made on a chip:
  - The cost to make each transistor reduces.
  - **But the chance that the chip will not work due to a defect rises.**

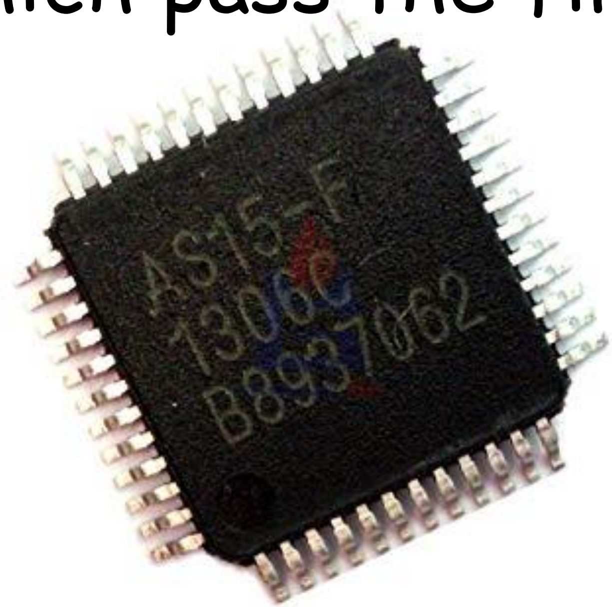
- Moore speculated 1965 that there is a transistor density or complexity:
  - At which "the minimum cost" is achieved.

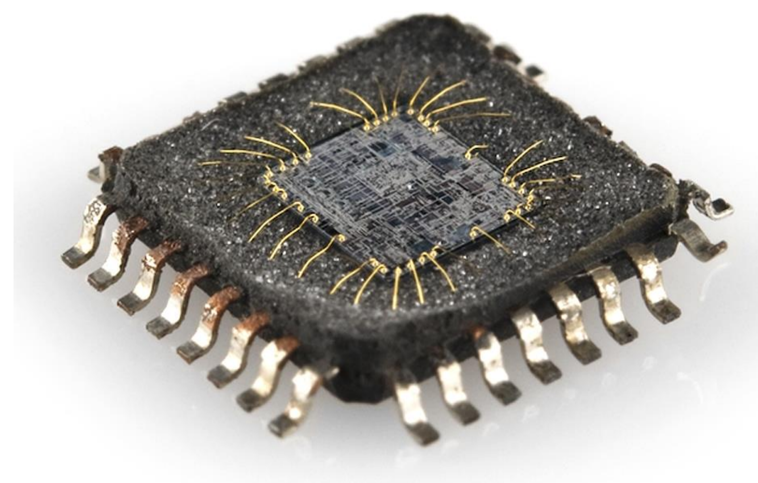


Integrated Circuits: Costs

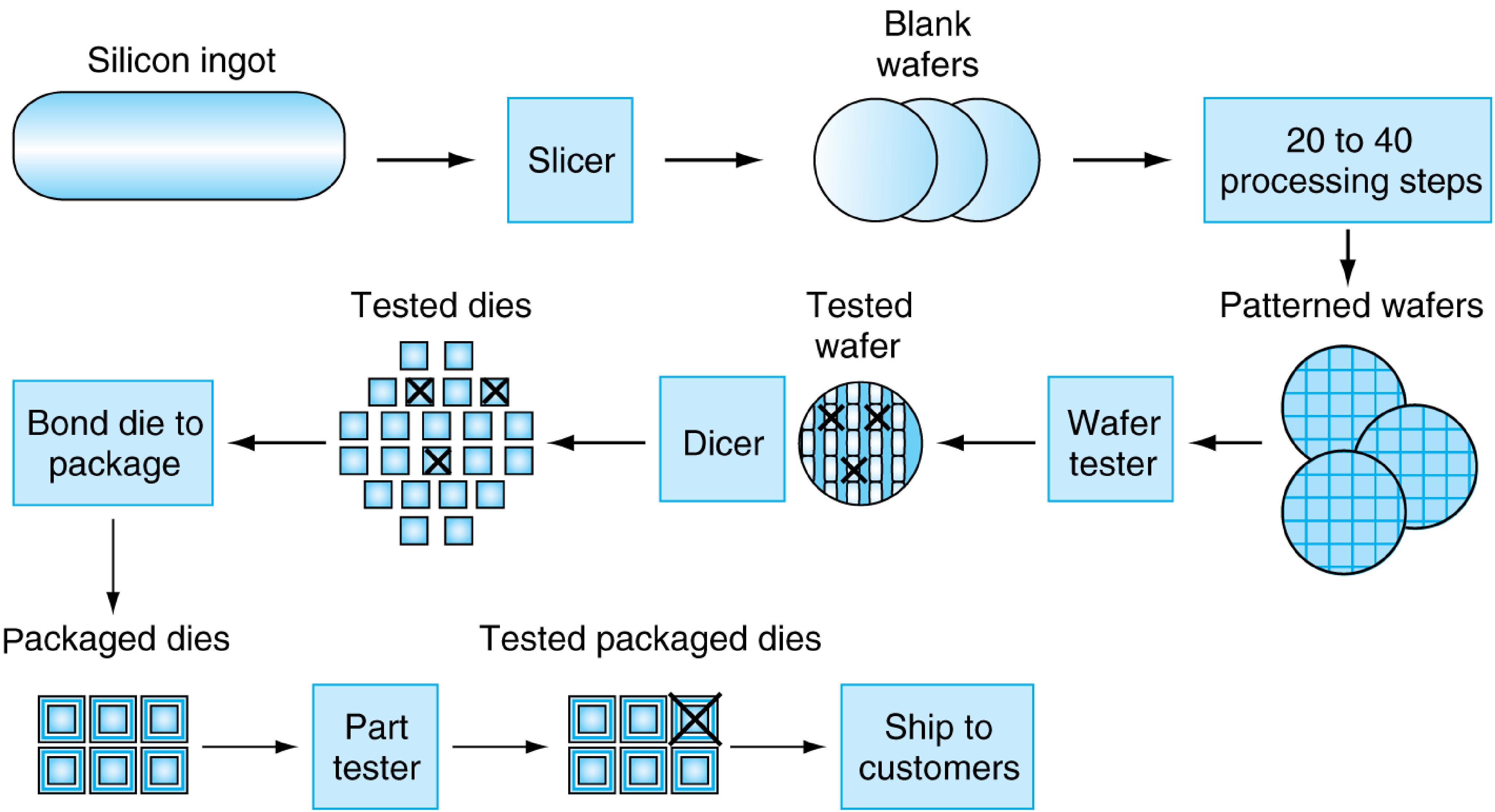
$$\text{IC cost} = \frac{\text{Die cost} + \text{Testing cost} + \text{Packaging cost}}{\text{Final test yield}}$$

Final test yield: Fraction of packaged dies which pass the final testing state.





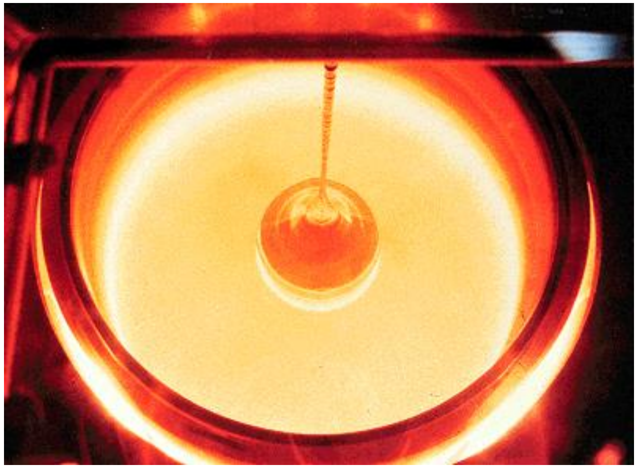
IC Manufacturing Process



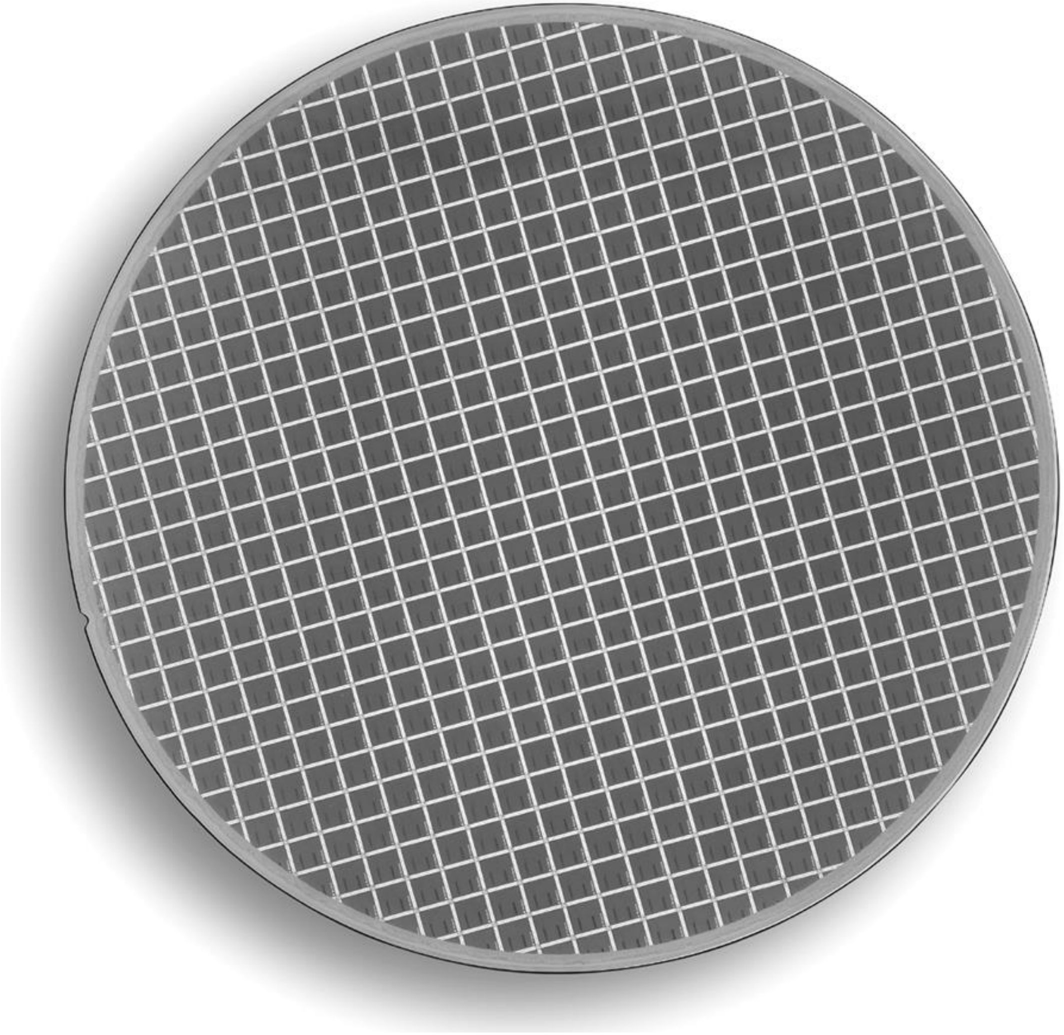
- Yield: proportion of working dies per wafer

8” MIPS64 Wafer (564 Dies)

Drawing single-crystal
Si ingot from furnace....



Then, slice into wafers and pattern it…



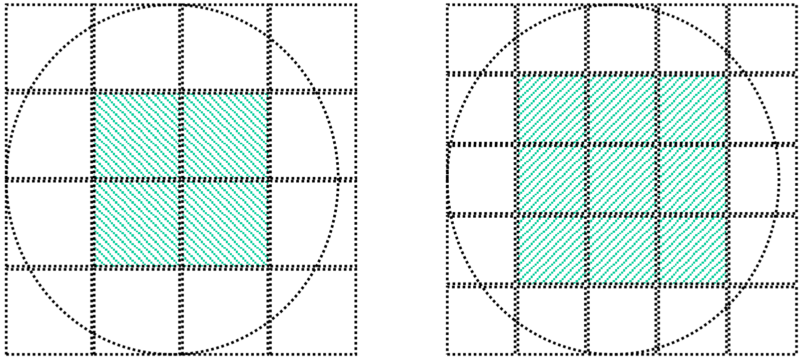
Integrated Circuits Costs

$$\text{IC cost} = \frac{\text{Die cost} + \text{Testing cost} + \text{Packaging cost}}{\text{Final test yield}}$$

$$\text{Die cost} = \frac{\text{Wafer cost}}{\text{Dies per Wafer} * \text{Die yield}}$$

Final test yield: Fraction of packaged dies which pass the final testing stage

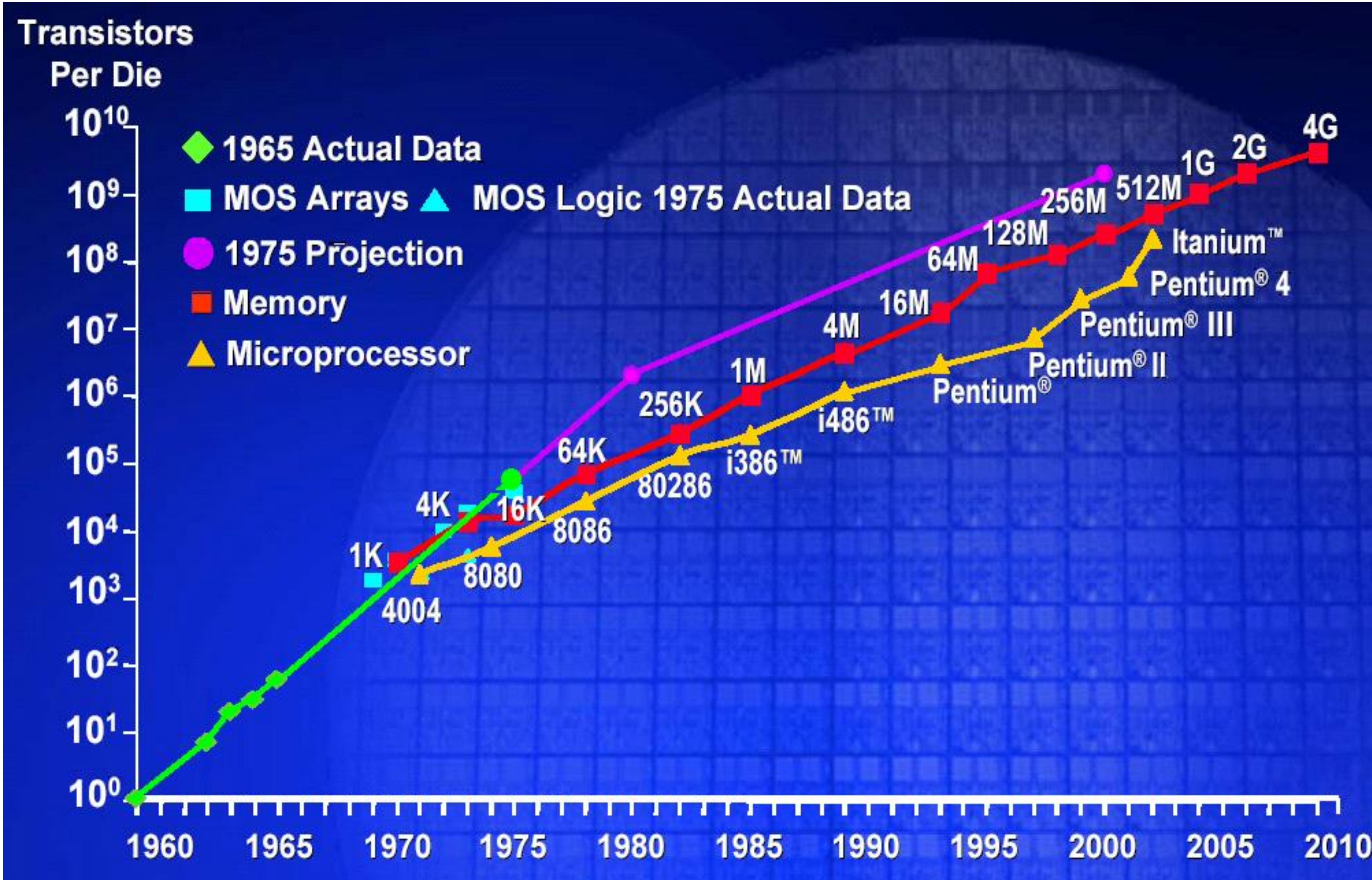
Die yield: Fraction of good dies on a wafer



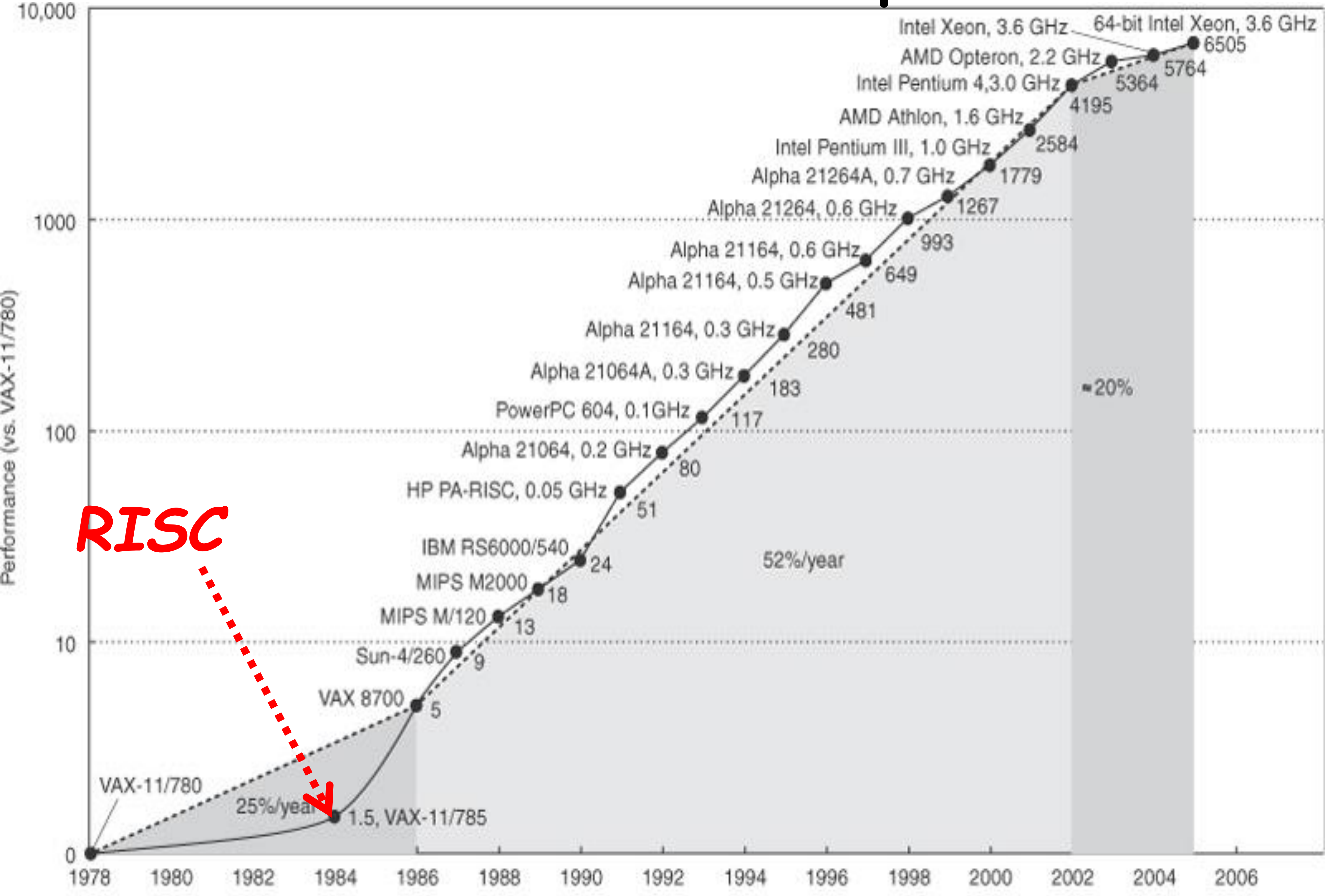
Mask Layers: Each IC is manufactured with 10 –15 mask layers.

- First half-dozen or so layers define transistors
- Other half-dozen or so define Interconnect

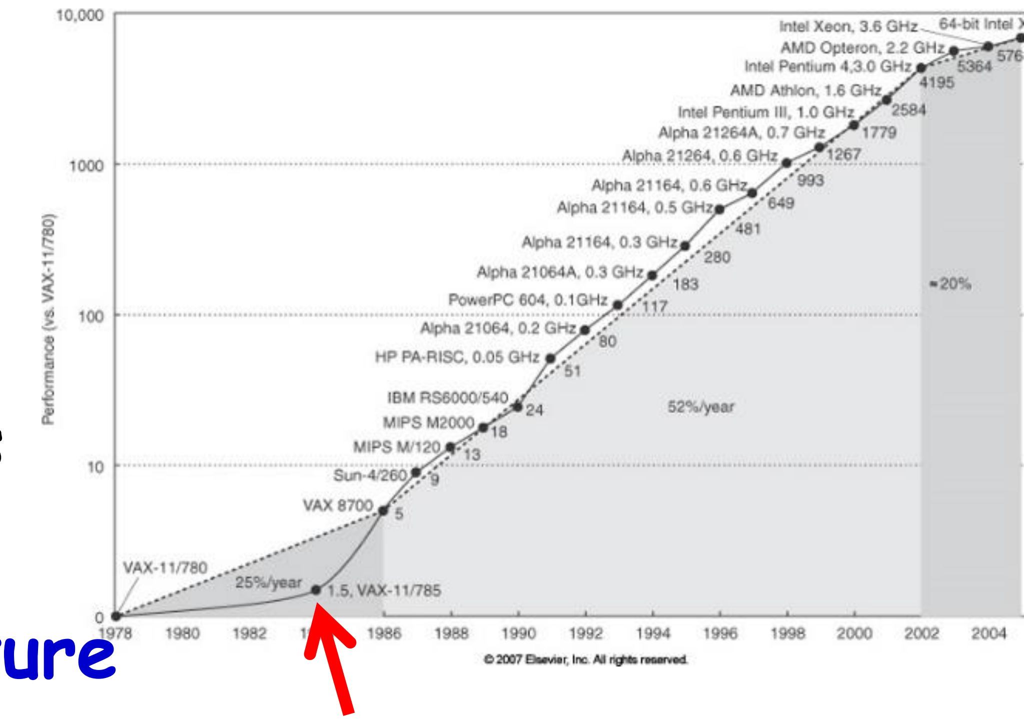
Number of Transistors in ICs



Processor Performance Improvements



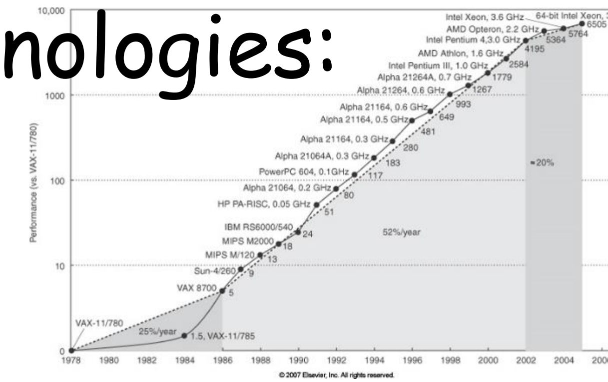
Where From Have the Performance Improvements Come?



- **Technology**
  - More transistors per chip
  - Faster logic
- **Processor Organization**
  - Effective and deep pipelines
  - Parallel execution units
- **Instruction Set Architecture**
  - Reduced Instruction Set Computers (RISC)
  - Multimedia extensions
  - Explicit parallelism
- **Compiler technology**
  - Finding more parallelism in code
  - Greater levels of optimization

How Did Processor Performance Improve?

- Till 1980s, most performance improvements came from innovative manufacturing technologies:
  - VLSI
  - Reduction in feature size

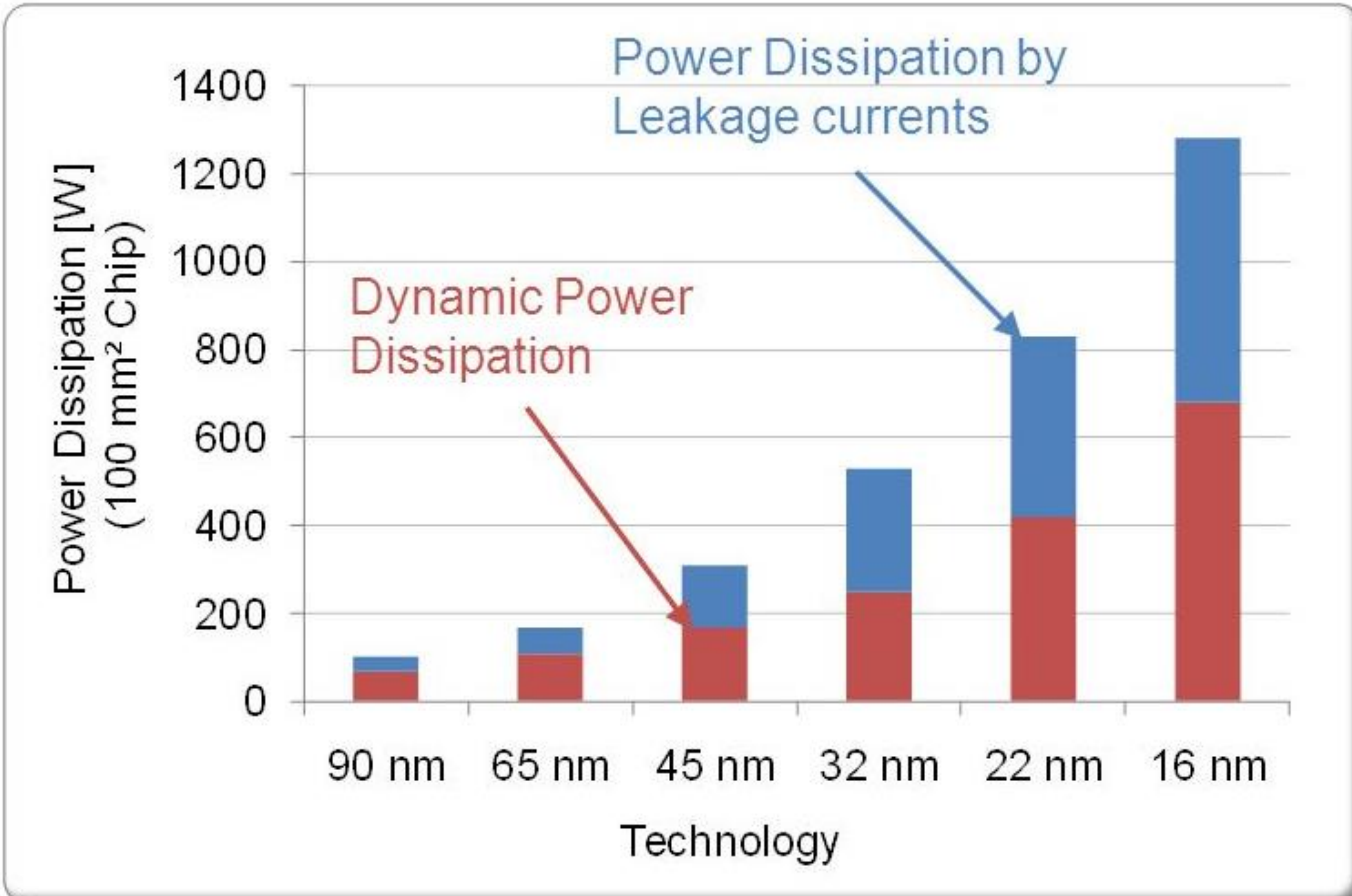


- **Improvements due to innovating manufacturing technologies have slowed down since 1980s:**
  - Smaller feature size gives rise to increased resistance, capacitance, propagation delays **and leakage**.
  - **Larger power dissipation.**

What is the power consumption of core i7 processor?

Roughly 100 watts when in use and 40 watts idle…

IC Power Consumption Trends



Source: S. Borkar (Intel),

Power Consumption in a Processor

- Power=Dynamic power + Leakage power
- Dynamic power = Number of transistors x capacitance x voltage^2 x frequency
- Leakage power is rising and will soon match dynamic power.

	Pentium	P-Pro	P-II	P-III	P-4
Year	1993	95	97	99	2000
Transistors	3.1M	5.5M	7.5M	9.5M	42M
Clock Speed	60M	200M	300M	500M	1.5G