Lect 3

Exercise

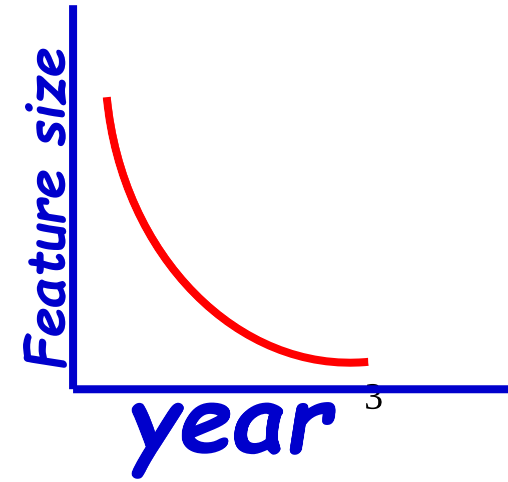
- Suppose by CPU redesign we managed to achieve:
  - 15% of capacitive load reduction
  - 15% voltage and 15% frequency reduction
  - **By how much will power consumption reduce?**

$$\frac{P_{new}}{P_{old}} = \frac{C_{old} \times 0.85 \times (V_{old} \times 0.85)^2 \times F_{old} \times 0.85}{C_{old} \times V_{old}^{\ 2} \times F_{old}} = 0.85^4 = 0.52$$

- ***The power wall***
  - ***We can't reduce voltage further***
  - ***We can't remove more heat***
- ***How else can we improve performance?***

Current Chip Manufacturing Process

- In early 2000 desktop processors were fabricated using a 65 nm process.
- Intel in 2007 began mass-producing 45nm chip (Nehalem) atom and core2duo processors.
- Intel in 2010 began mass-producing 32nm chip core i7 (Westmere architecture).
- Intel in 2017 began mass-producing 14nm chips
  - Compare: The diameter of an atom is of the order of 0.1 nm.
- A decade ago, chips were built using a 500 nm (0.5 micron) process.
- In 1971, 10micron process was used.



Manufacturing Process

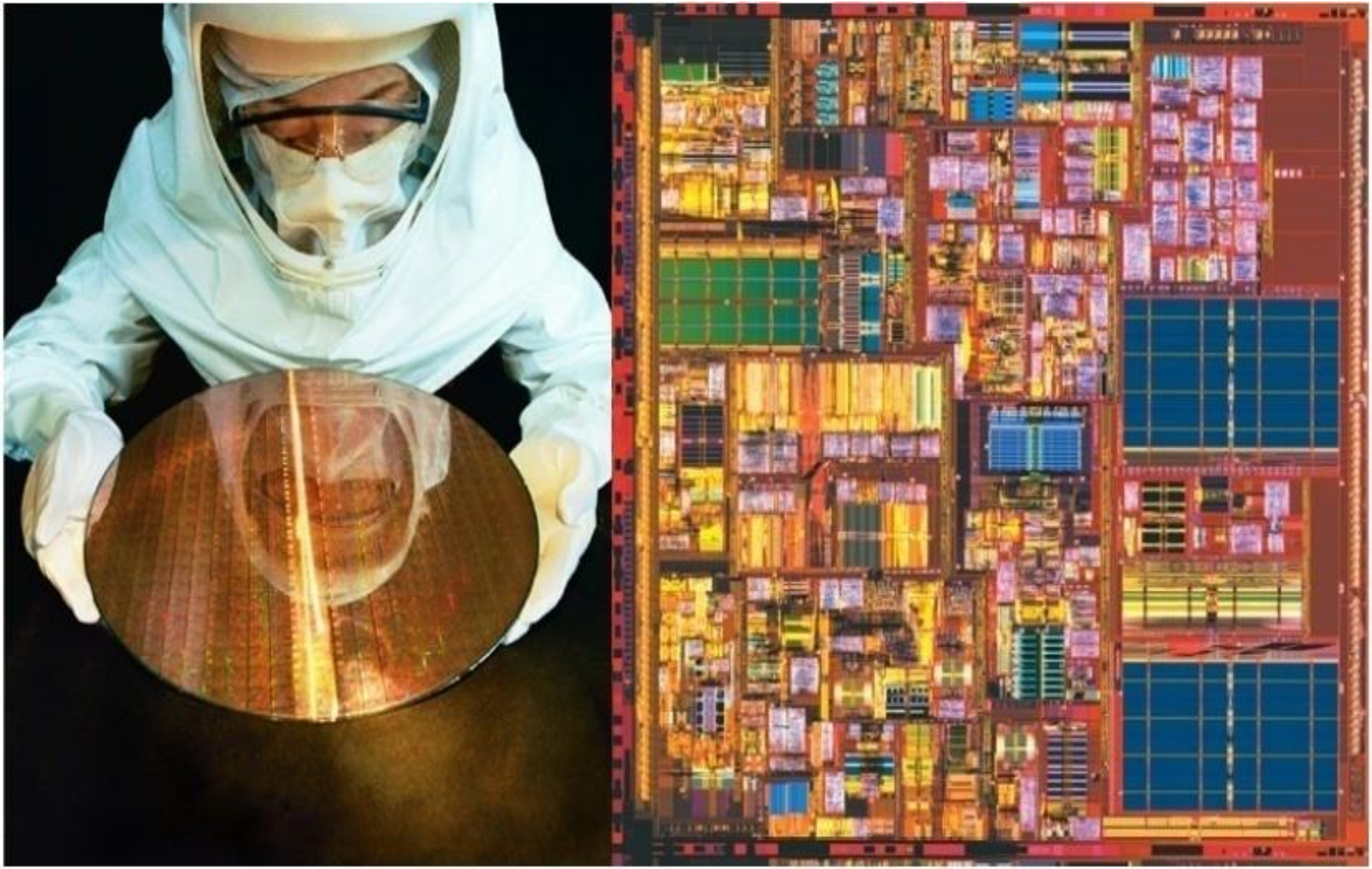
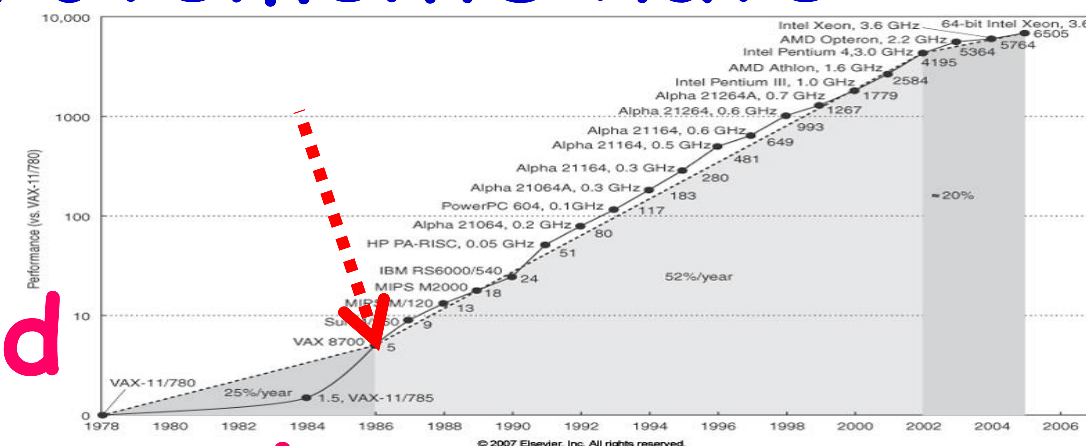


Photo courtesy of Intel

How Did Performance Improve? Cont…

- Since 1980s, large part of the performance improvements have come from:
  - Architectural and organizational innovations…



- What is the difference between:
  - Computer architecture and computer organization?

Architecture vs. Organization

- **Architecture:**
  - Also known as Instruction Set Architecture (ISA)
  - Programmer visible part of a processor: instruction set, registers, addressing modes, subroutine calls, etc.
- **Organization:**
  - High-level design: How many caches? How many arithmetic and logic units? What type of pipelining, control design, etc.
  - **Sometimes known as micro-architecture.**

Architecture

"… the attributes of a processor as seen by the programmer, i.e., the conceptual structure and functional behavior, as distinct from the organization of the data flows and controls, the logic design, and the physical implementation."

in **Architecture of the IBM system/360**, Amdahl, Blaauw, and Brooks, 1964

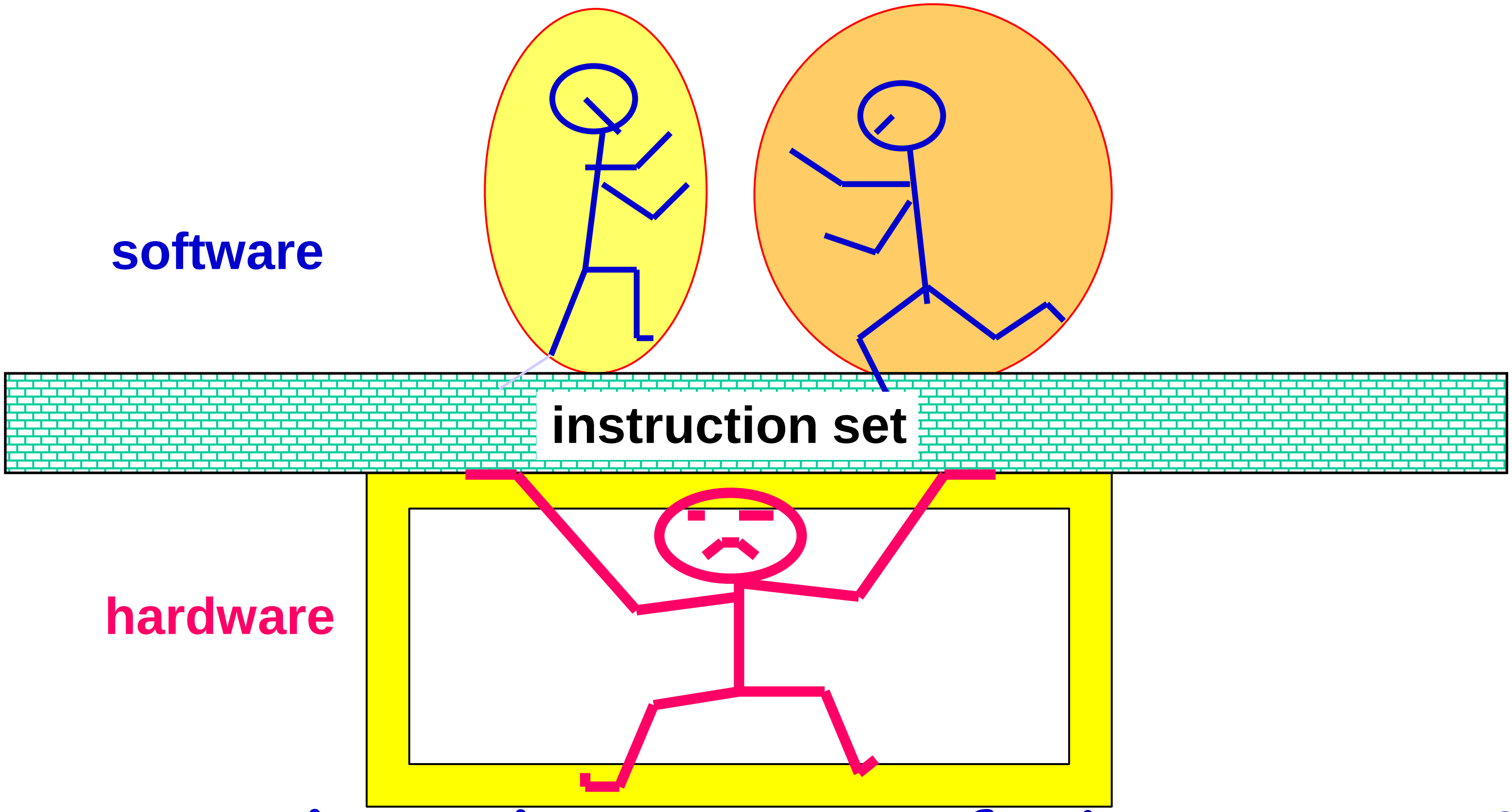
Instruction Set Architecture (ISA)

- ISA defines the aspects of a computer that an assembly language programmer must understand:
  - To be able to write a correct program for that machine.
- A family of computers having the same architecture should be able to run the same assembly language program.
  - **Architecture leads to the notion of binary compatibility.**

Instruction Set Architecture (ISA)

- Defines aspects of a computer:
  - An assembly language programmer must understand ISA to be able to write correct programs for that machine.
- A family of computers having the same architecture should be able to run the same assembly language program.
  - **Architecture leads to the notion of binary compatibility.**

Instruction Set Architecture: Abstracts Hardware



- **What are the advantages of abstraction?**
  - Lasts through many generations (portability)
  - Used in many different processors (generality)
  - Provides convenient functionality to higher levels
  - Permits efficient implementation at lower levels

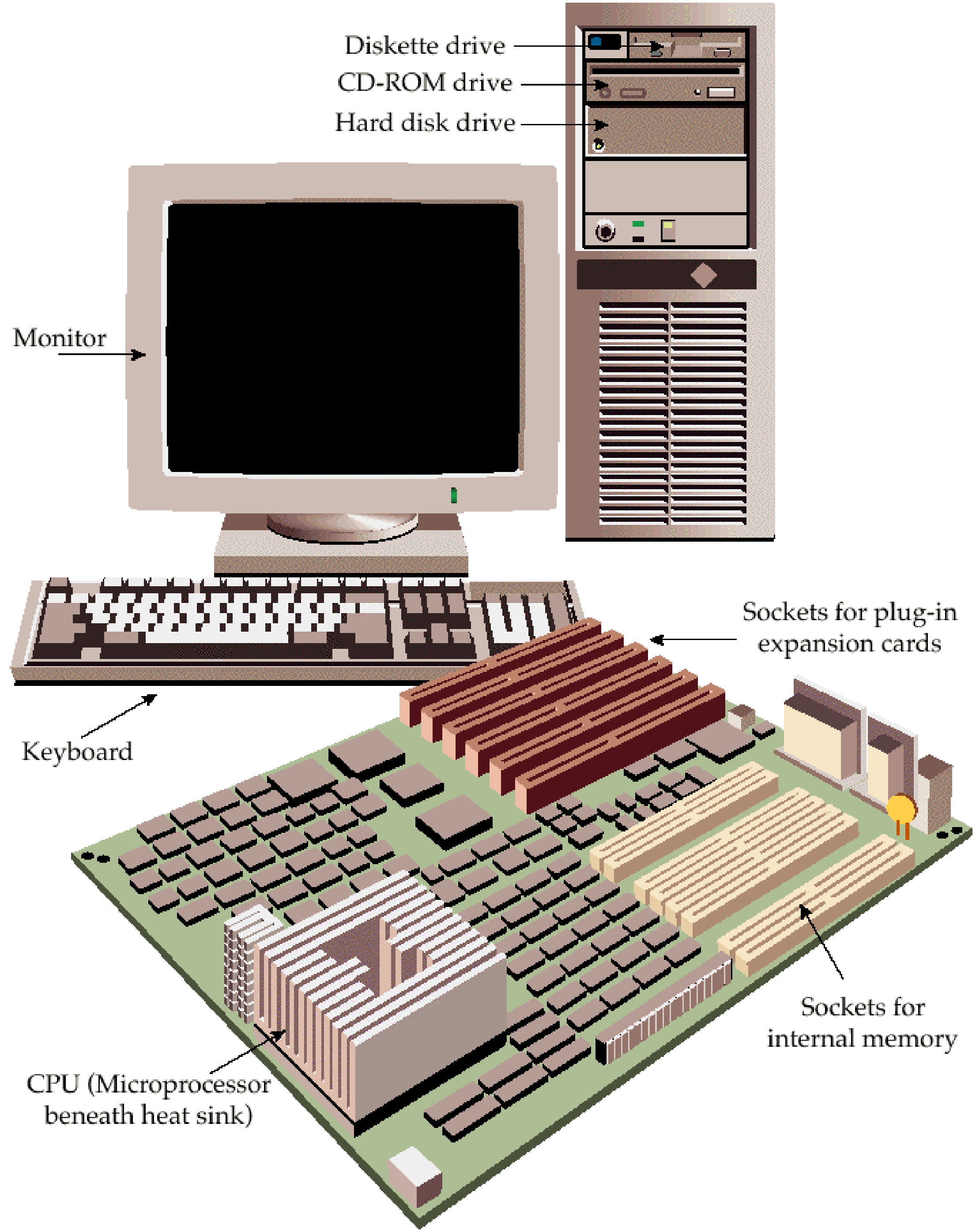
Present Computer Architecture

Computer Architecture loosely being used to cover all three aspects of computer design.

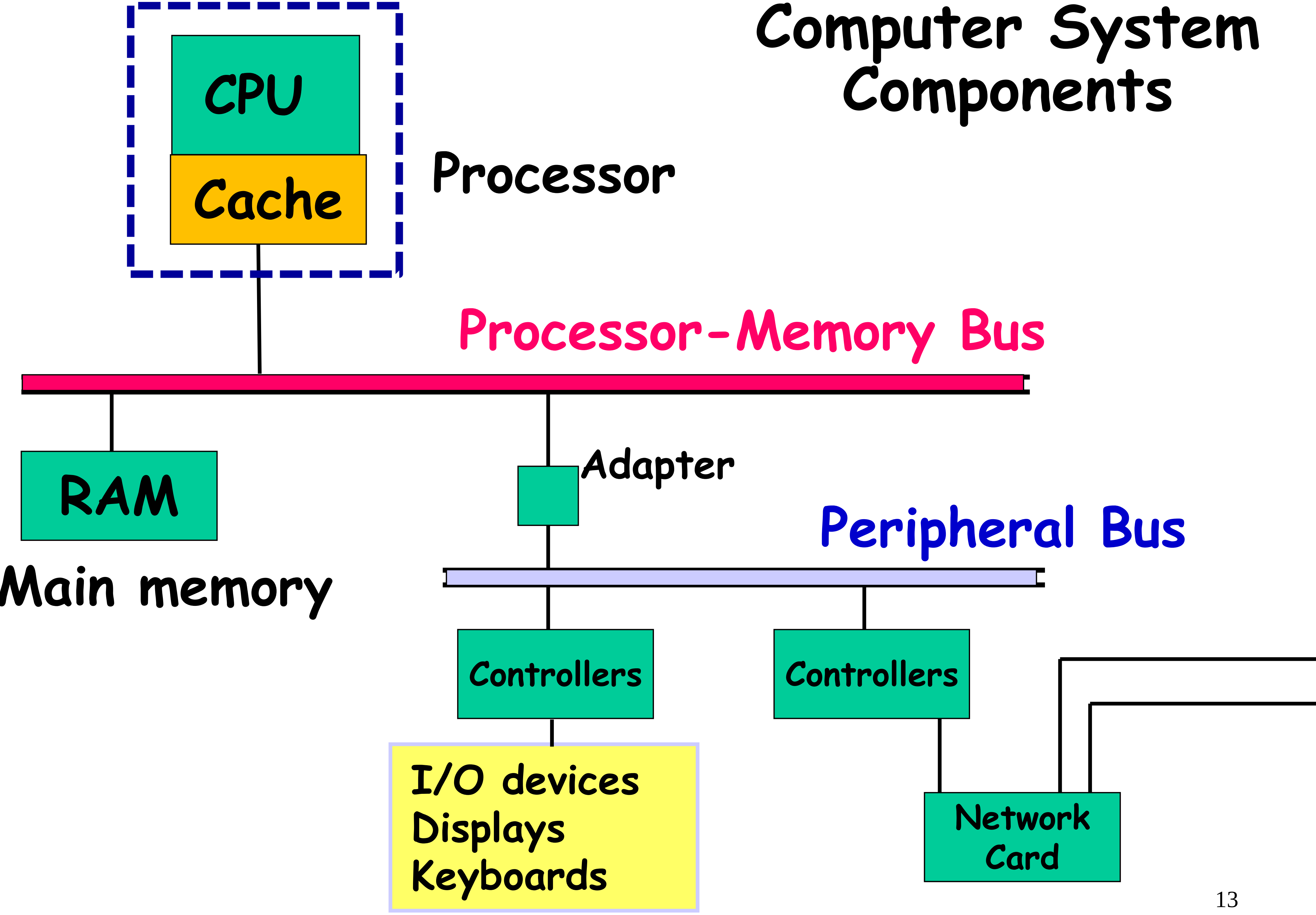
- **Instruction Set Architecture**
  - The computer as visible to the assembly language programmer or compiler writer (registers, data types, instruction set, instruction formats, addressing modes)
- **Organization**
  - High level aspects of computer's design such as the memory system, the bus structure, and the internal CPU (datapath + control) design
- **Hardware**
  - Detailed logic design, interconnections, external connections

A Typical Computer System

Exercise: Draw a schematic diagram to show how the different components of a computer are interconnected by bus(es).



Computer System Components



Diagnostic Quiz

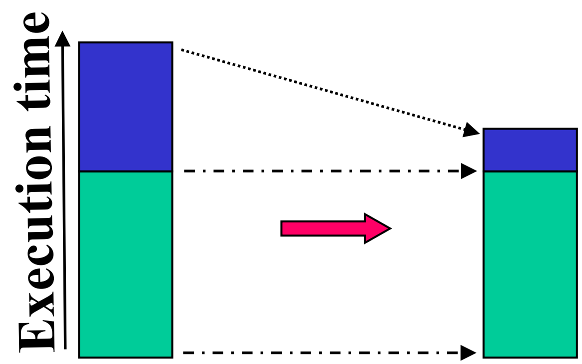
- Consider a program in which memory access constitutes 40% of the execution time.
  - We replace the RAM of the computer with a faster one and now memory access is 10 times faster.
- What would be the Speedup in program execution?

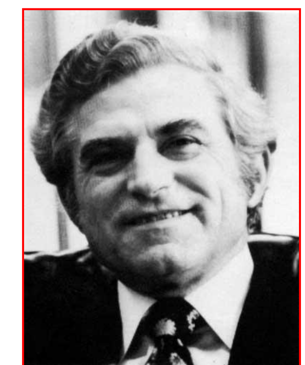
- Fraction enhanced = 0.40
- Performance enhanced = 10

$$Speedup = \frac{1}{1 - 0.4 + \frac{0.4}{10}} = \frac{1}{0.64} \approx 1.56$$

Amdahl's Law

- Quantifies overall performance gain due to an improvement to some part of a computation.





Amdahl

- Amdahl's Law:
  - The speedup achieved from using some faster mode of execution is limited by the amount of time that enhancement is actually used.

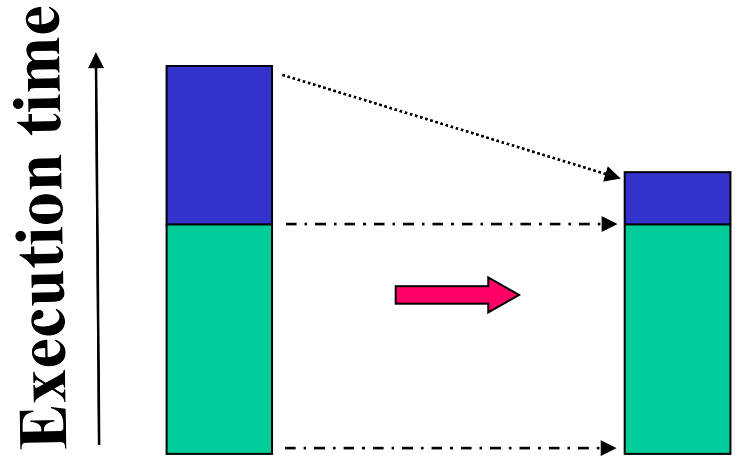
$$\text{Speedup} = \frac{\text{Execution time for a task without enhancement}}{\text{Execution time for the task using enhancement}}$$

Amdahl's Law Equations

$$\text{Execution time}_{new} = \text{Execution time}_{old} \times \left[(1 - \text{Fraction}_{enhanced}) + \frac{\text{Fraction}_{enhanced}}{\text{Speedup}_{enhanced}} \right]$$

$$\text{Speedup}_{overall} = \frac{\text{Execution Time}_{old}}{\text{Execution Time}_{new}} = \frac{1}{(1 - \text{Fraction}_{enhanced}) + \frac{\text{Fraction}_{enhanced}}{\text{Speedup}_{enhanced}}}$$

Use previous equation,
Solve for speedup



Don't just try to memorize
these equations and plug numbers into them.
It's always important to think about the problem too!

Exercise 1: Answer

- Floating point instructions improved to run 2X faster.
  - But only 10% of execution time spent on FP execution.

$$ExTime_{new} = ExTime_{old} \times (0.9 + 0.1/2) = \boxed{0.95 \times ExTime_{old}}$$

$$Speedup_{overall} = \frac{1}{0.95} = 1.053$$

$$\boxed{= 5.3\%}$$

Amdahl's Law: Exercise 1

- Consider that the floating point (FP) unit of a computer was improved to run 2 times faster.
- However, only 10% of execution time is spent on FP instructions

$\text{ExTime}_{new} = ?$

$\text{Speedup}_{overall} = ?$