CSE 360-Computer Architecture

Lecture-1

Dr. Shamim Akhter

Associate Professor

Department of Computer Science and Engineering

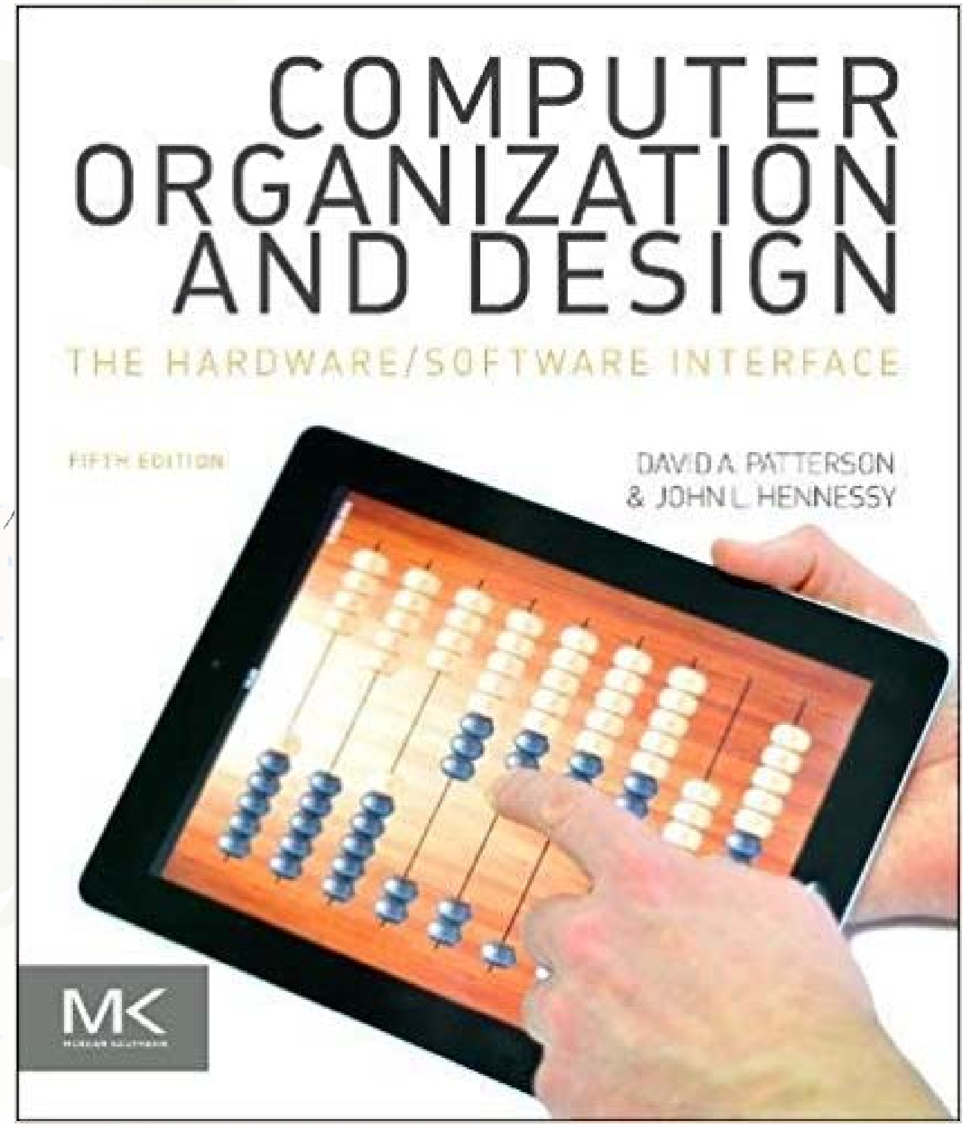
email: shamimakhter@ewubd.edu

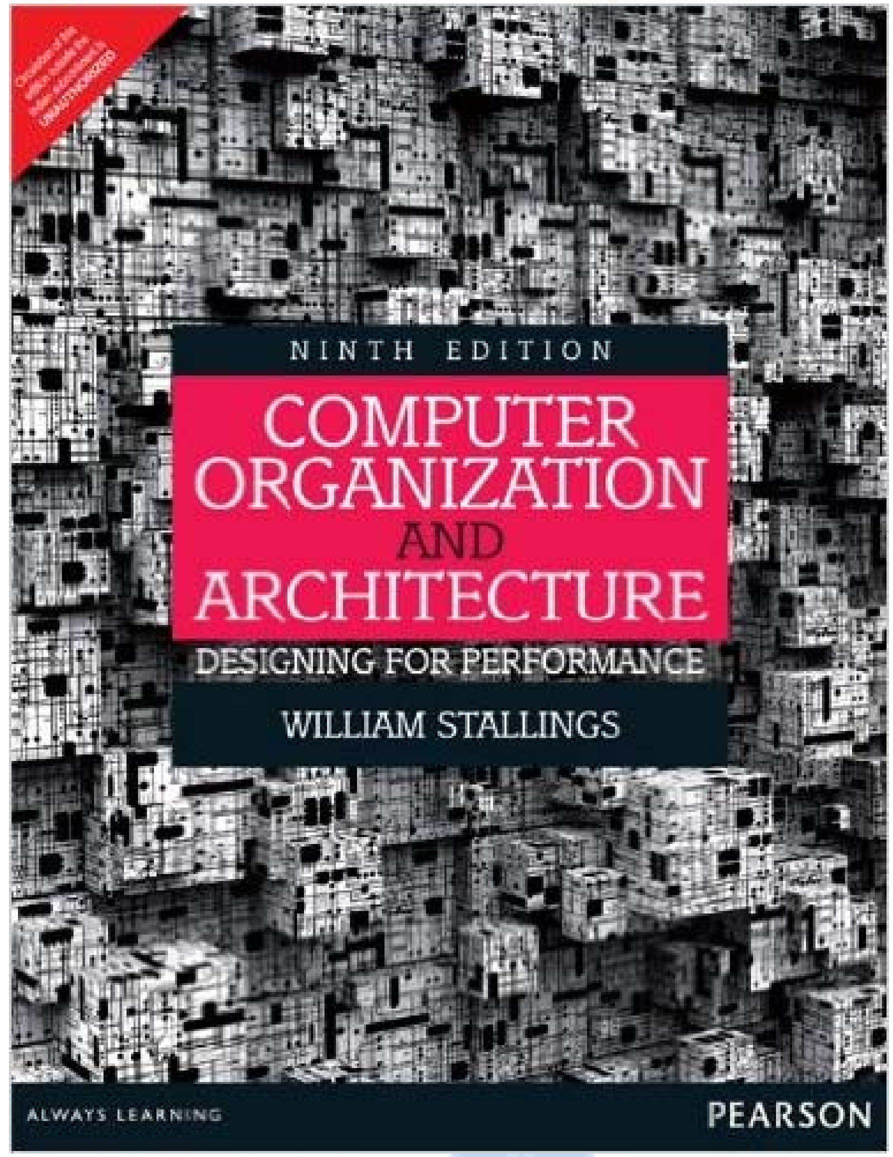
WELCOME TO

CSE 360-Computer Architecture

Spring, 2018

Textbooks


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


NINTH EDITION
COMPUTER ORGANIZATION AND ARCHITECTURE
DESIGNING FOR PERFORMANCE
WILLIAM STALLINGS
ALWAYS LEARNING
PEARSON

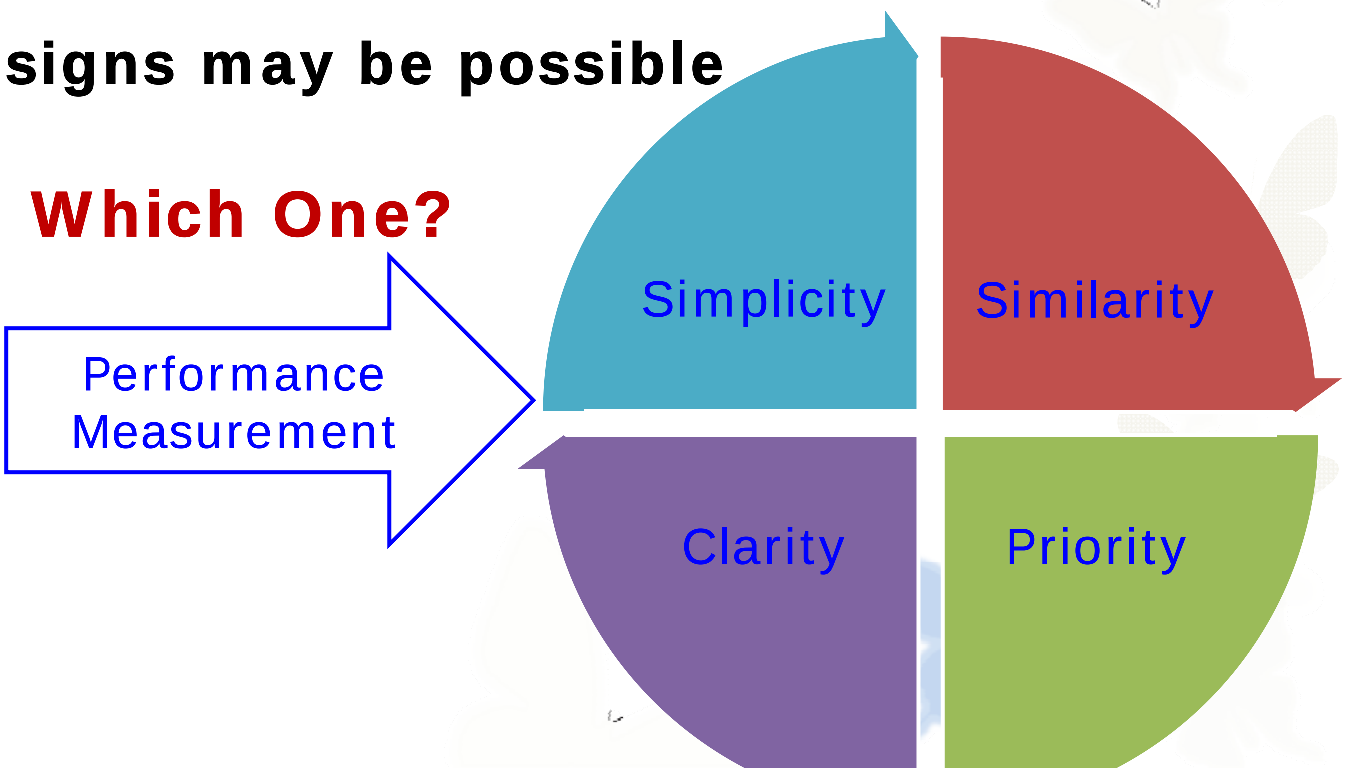
Course Major Objective-

To Learn/Understand Computer System with relationship between H/W and S/W

- Simple understand how computer works?
  - What is good design and what is not?

Several designs may be possible

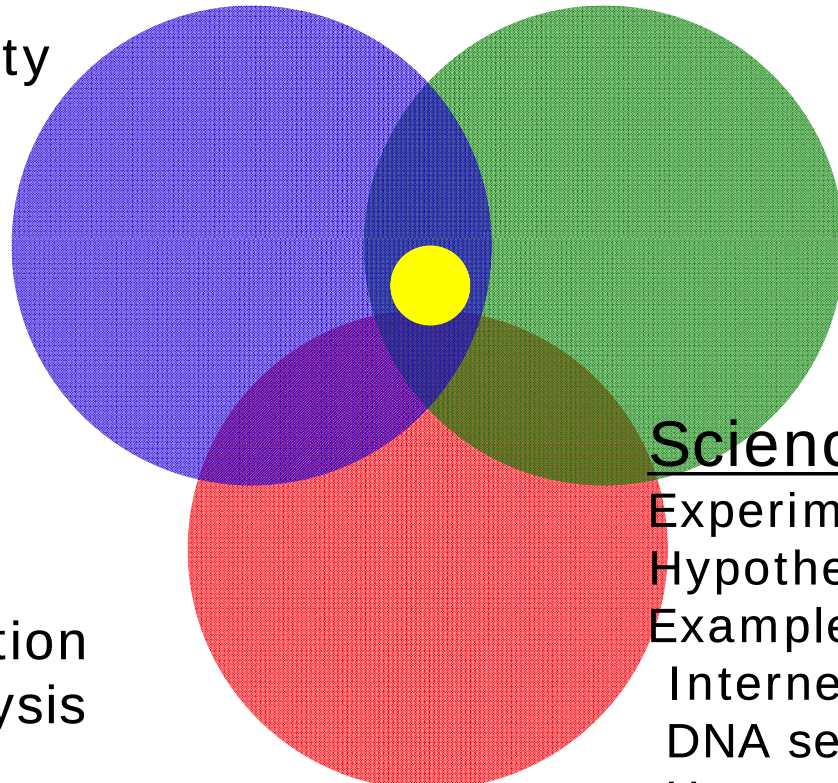
Which One?



CSE is an Estuary

Where does architecture fit into computer science?
Engineering? Or some Science?

<u>Engineering</u>
Design
Handling complexity
Real-world impact
Examples:
Microprocessor



<u>Science</u>
Experiments
Hypothesis
Examples:
Internet behavior,
DNA sequence supercomputer
Human/computer interaction

<u>Mathematics</u>
Limits of computation
Algorithms & analysis
Cryptography
Logic
Proofs of correctness

<u>Other Issues</u>
Public policy, ethics,
law, security

Text Books have a common name:

a) Computer Organization

b) Design / Architecture

Interesting !!

What is Computer Architecture?

Computer Architecture is the science and art of

-selecting and interconnecting hardware components

WHY?

To Create Computers

How?

To Meet Some Goals

Functional

Performance

Cost

- Analogy to architecture of buildings…

What is ~~Computer~~ Architecture

- The role of a building architect:

Materials
Steel
Concrete
Brick
Wood
Glass

Design

Plan

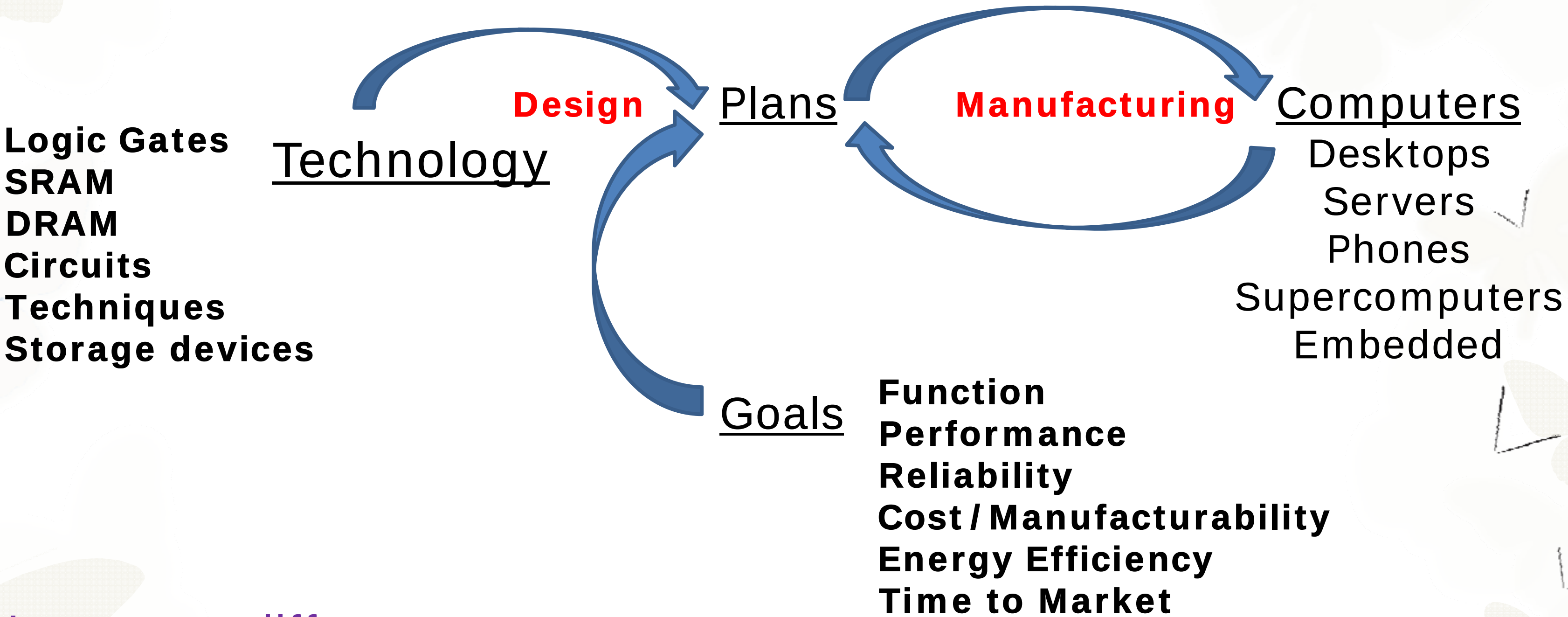
Construction

Buildings
Houses
Offices
Apartments
Stadiums

Goals
Function
Cost
Safety
Energy Efficient

What is Computer Architecture

- The role of a computer architect:



Important differences:

- age (~80 years vs thousands),
- rate of change (technology, applications, goals)

Rate of Changes

Technology

[Steve Jobs](#)
The Man in The Machine

https://www.slideshare.net/BSGAfrica/top-5-considerations-when-choosing-and-managing-a-technology-vendor

Color
Thinner
Faster Processor
New version OS –IOS
Camera resolution
Screen resolution

“Technology”

Year	Technology used in computers	Relative performance/unit cost
1951	Vacuum tube	1
1965	Transistor	35
1975	Integrated circuit	900
1995	Very large scale integrated circuit	2,400,000
2005	Ultra large scale integrated circuit	6,200,000,000

ON/OFF switches-control electricity

100 transistors into a single chip

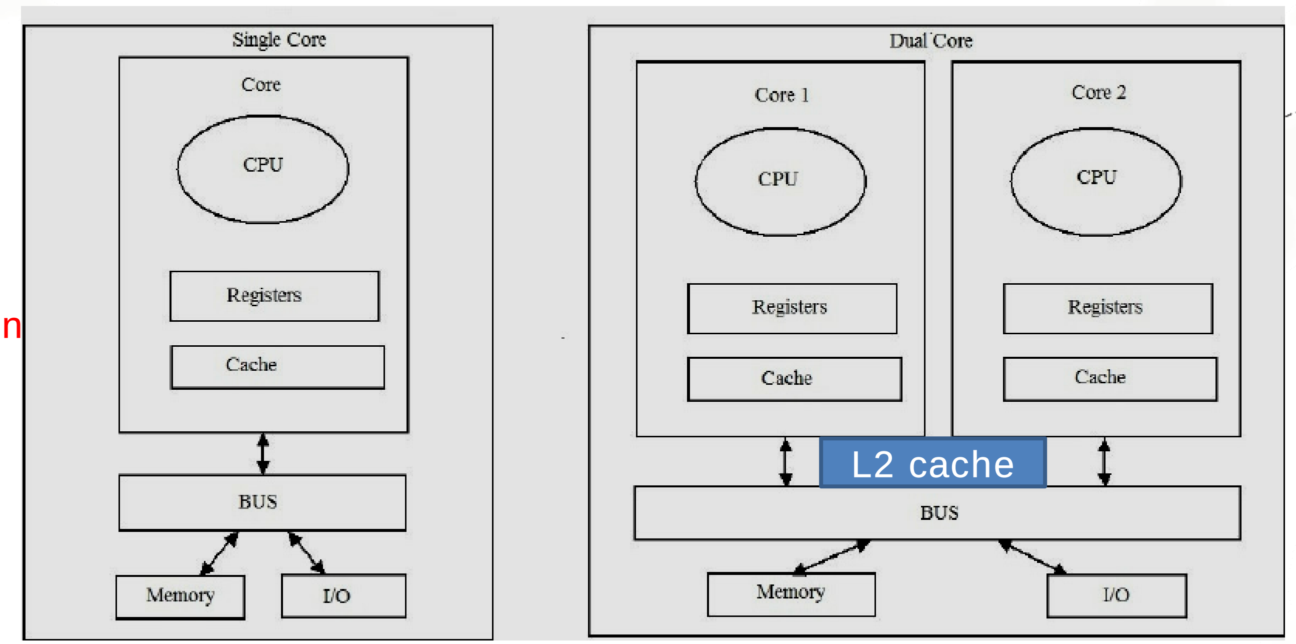
Microprocessor- VLSI device

High Performance, High Reliability

- Several kinds of integrated circuit families
  - SRAM: optimized for speed (used for processors)
  - DRAM: optimized for density, cost, power (used for memory)
  - Flash: optimized for density, cost (used for storage)
- Increasing opportunities for integrating multiple technologies
  - Inter-connection technologies
  - Disk, optical storage, ethernet, fiber optics, wireless

Interconnection technology: Example

One instruction



Multiple instruction execution

Single Processor	Multi Processors SMP	Multicore CMP	MIC/ Many integrated cores 50 cores in a die	Simultaneous Multithreading/ SMT
		Homogeneous collection of general-purpose processors on a single chip		Core's components are duplicated
		GPU – graphics, video etc processing core Plug-in graphics card (display adapter)		Pipeline is capable to run multiple instructions from multi-threads

Interconnection technology: Example

We want more! **100 cores in a single die !**

- Need to make smaller cores but lose the functionality
- Making a bigger die but increases cost

What else???

How about to reduce the cost of **chip-to-chip communication**
- power, bandwidth, latency

Break a multi-core chip into a many-chip-system
-smaller chips lead to higher yields and lower cost
-different chips lead to system adaptability and reconfigure ability
-aggregate systems of chips effectively

Interconnection
Technology
Exploration

How does the interconnection technology change the word?

Example: Wire Technology

1. Alexander Graham Bell invented the telephone in 1876,messages were traveled as electric currents and transmitted over copper wire.

2. Need better sound quality, cover greater distances, greater capacity
   - integration of metallic two-wire circuits, loading coils, vacuum-tube amplifiers, coaxial cable, and microwave radio relay systems.

3. Then came conversion from Analog Signal to Digital Signal
   - achieved more frequency, greater capacity
   - use in TV and Digital Computer

4. Need to carry information much faster
   - Solution Laser ->Optoelectronics

   - Transmission rate-10kb/S
   - Problem: Clouds, Haze, Rain ---> Block beam.

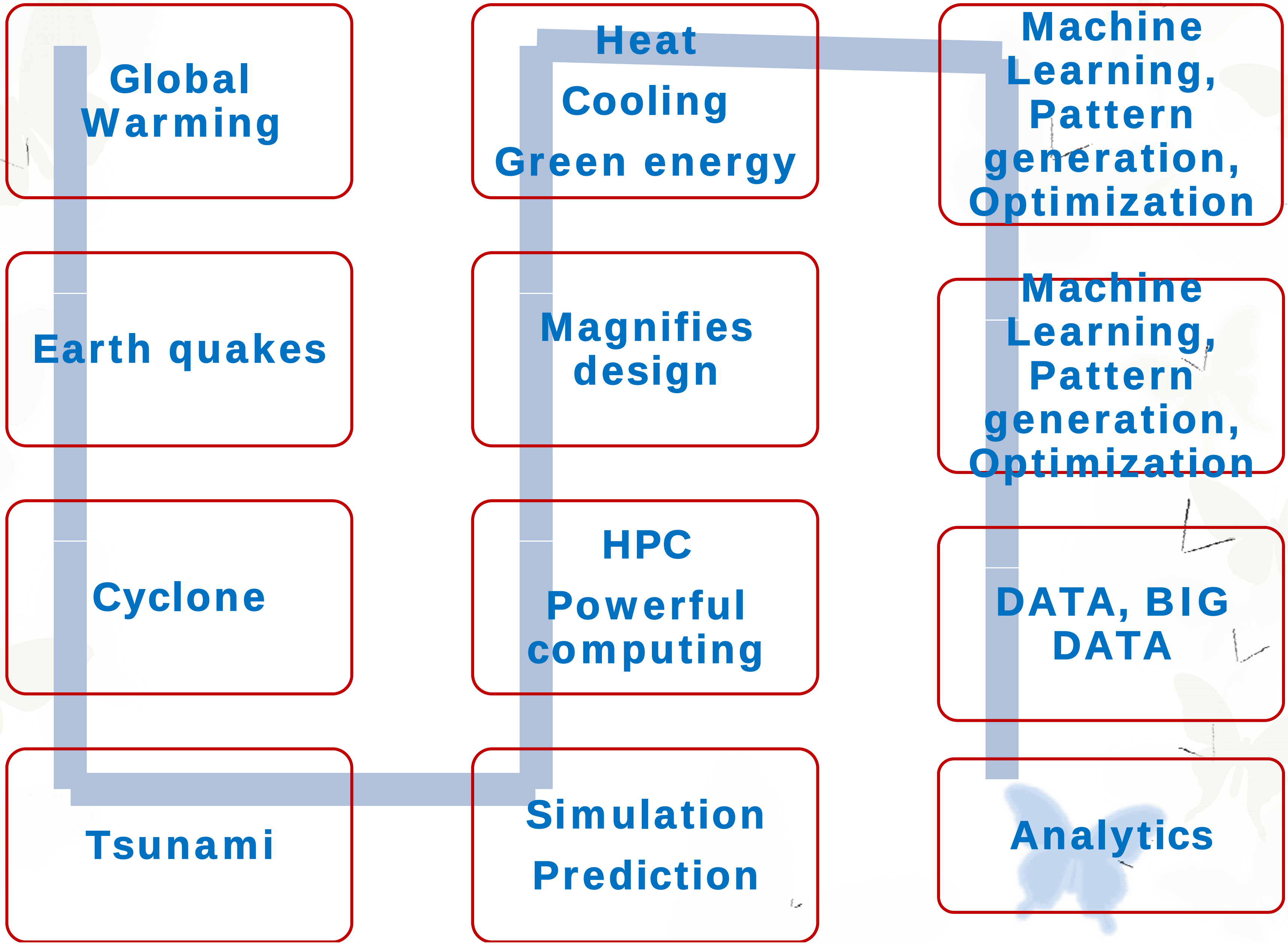
5. Use laser inside glass fibers
   -achieved 100Mb/S

6. Can data transmit as the speed of light?
   More higher frequency (Gb/S, Tera b/S .....)
   Replace microwave to light wave

   - Light waves ---->
   Noble Prize in Physics-2009, Prof. Charles K. Kao

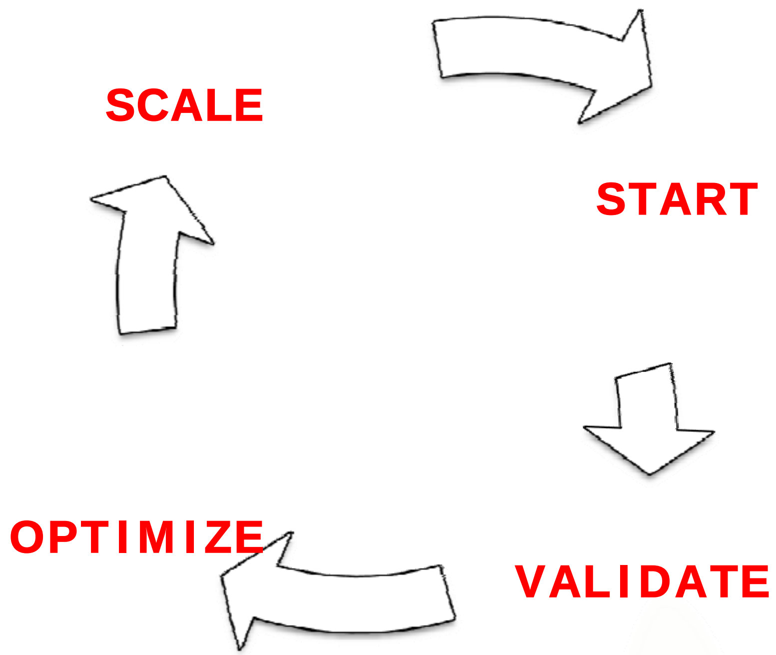
Rate of Changes

Applications

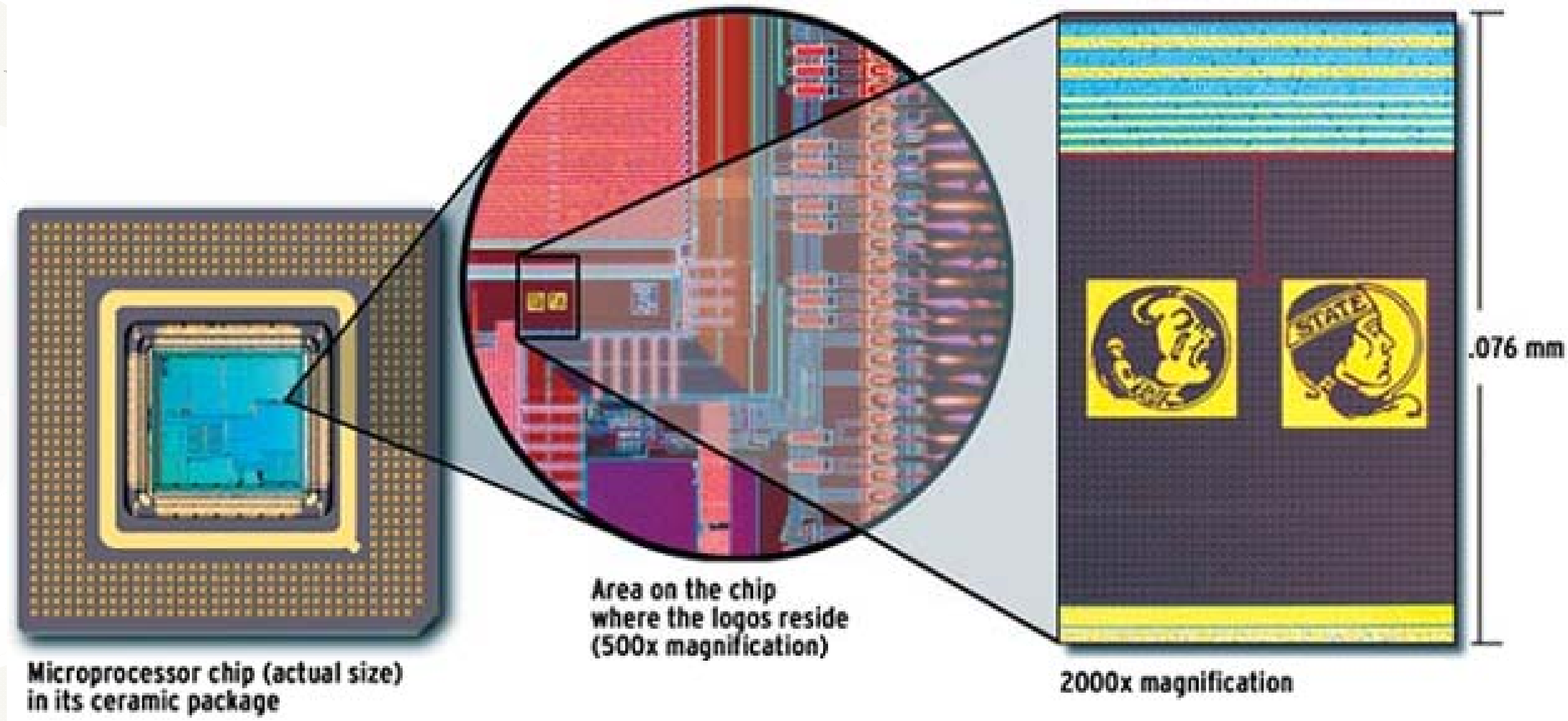


Boot-strapping Design

- Building a system using itself/ predecessor ver.



Automated Mass Production
(magnifies design-over millions of chips)



Source: https://spectrum.ieee.org/image/MTQxNTQ3Nw

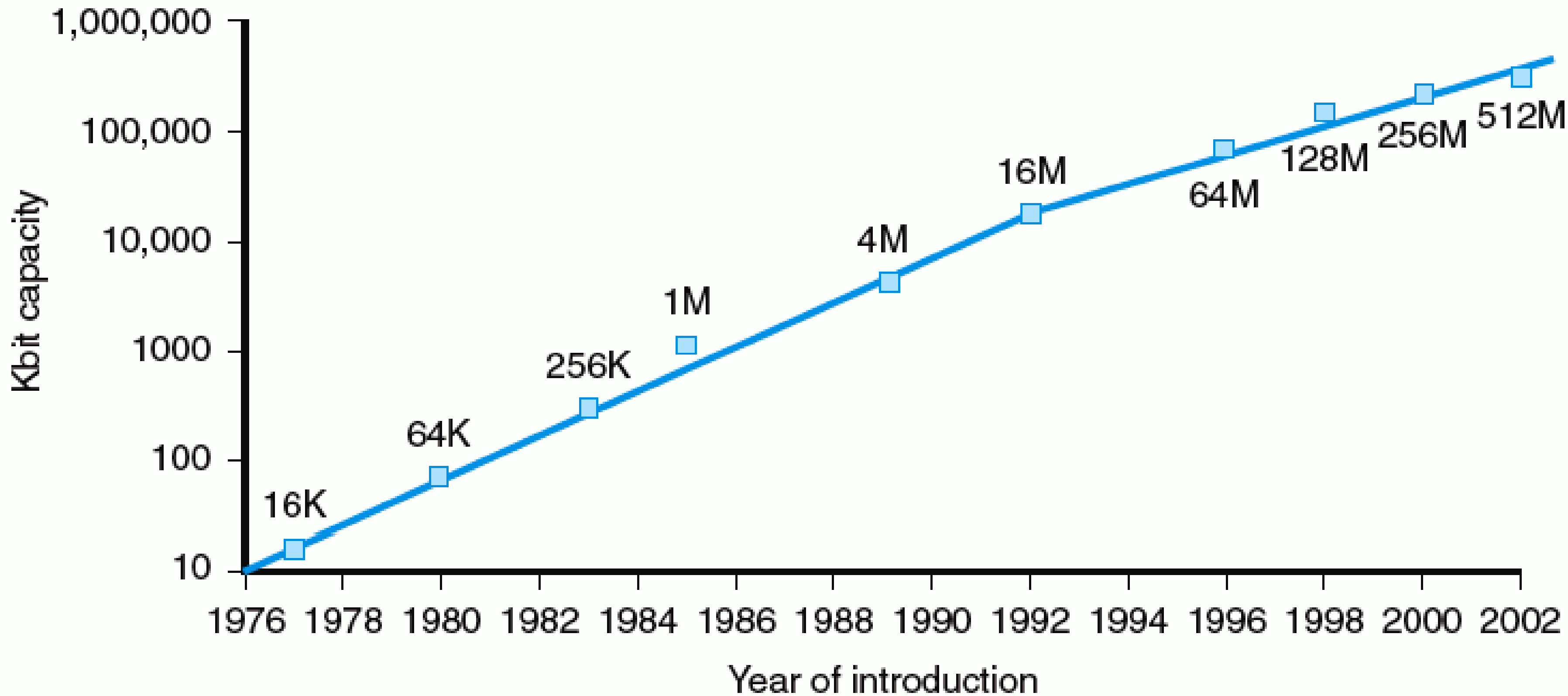
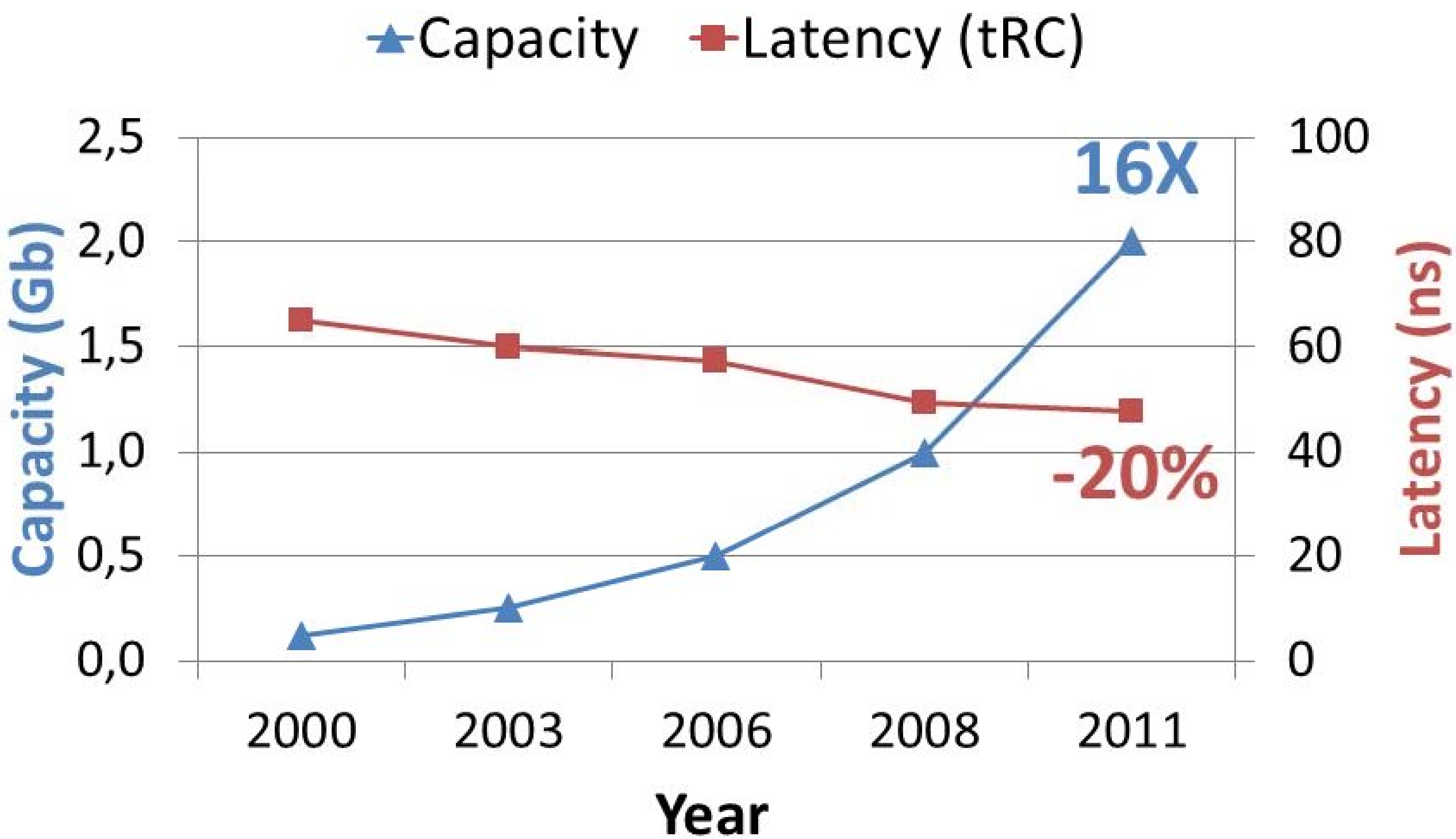
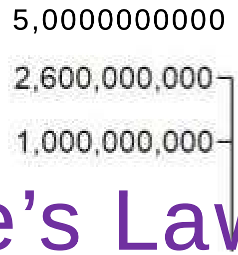


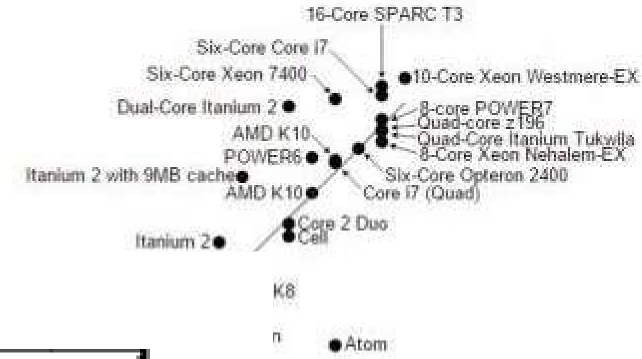
FIGURE 1.13 Growth of capacity per DRAM chip over time. The *y*-axis is measured in Kbits, where K = 1024 (2^{10}). The DRAM industry quadrupled capacity almost every 3 years, a 60% increase per year, for 20 years. This "four times every three years" estimate was called the *DRAM growth rule.* In recent years, the rate has slowed down somewhat and is somewhat closer to doubling every two years or four times every four years.

DRAM Latency-Capacity Trend



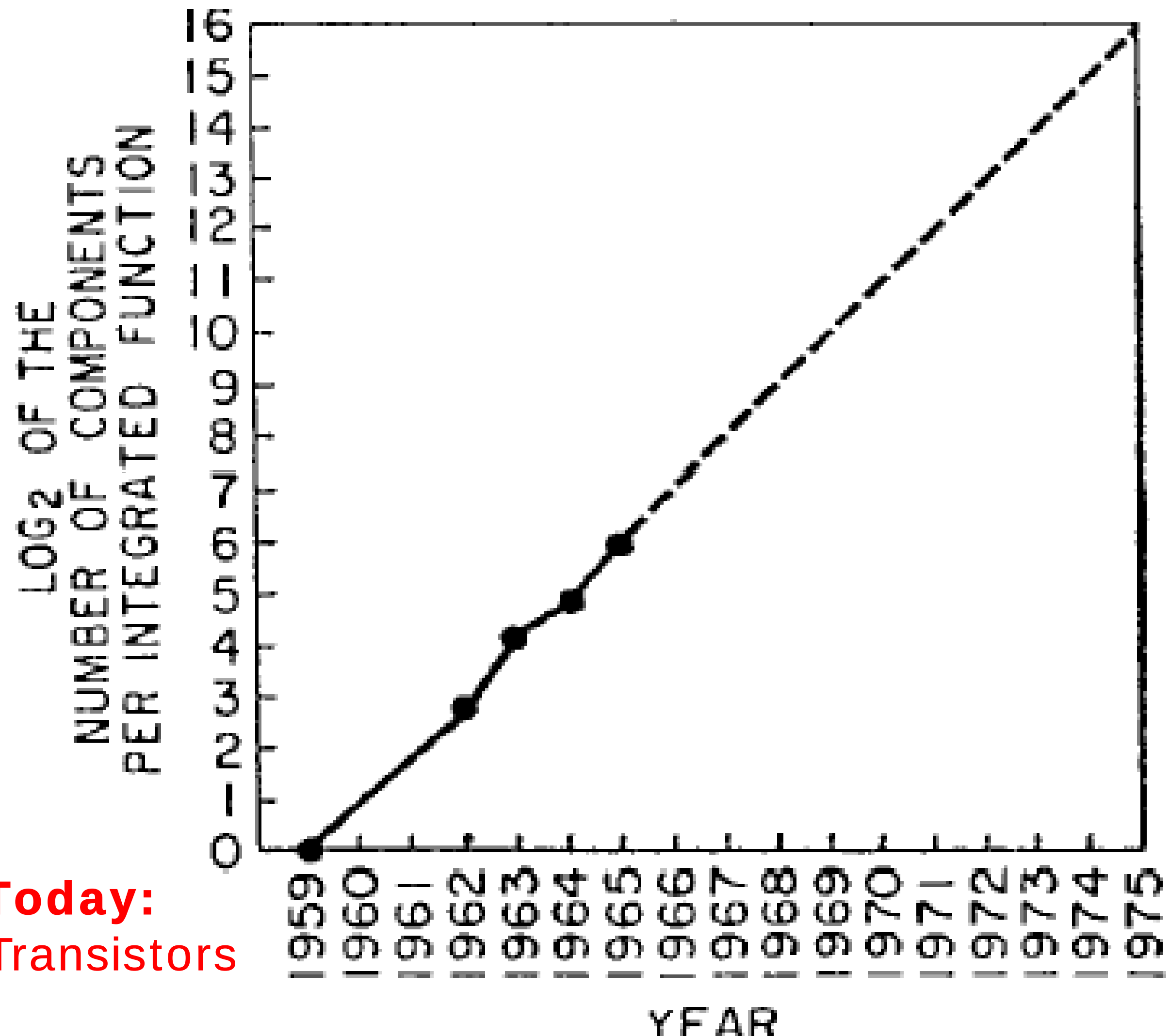
DRAM latency continues to be a critical bottleneck





2^x

Moore's Law - 1965



Transistor count doubling in every two years

Today:
2^{32} Transistors

@ Gordon E. Moore, "Cramming More Components onto Integrated Circuits," Electronics, pp. 114–117, April 19, 1965.

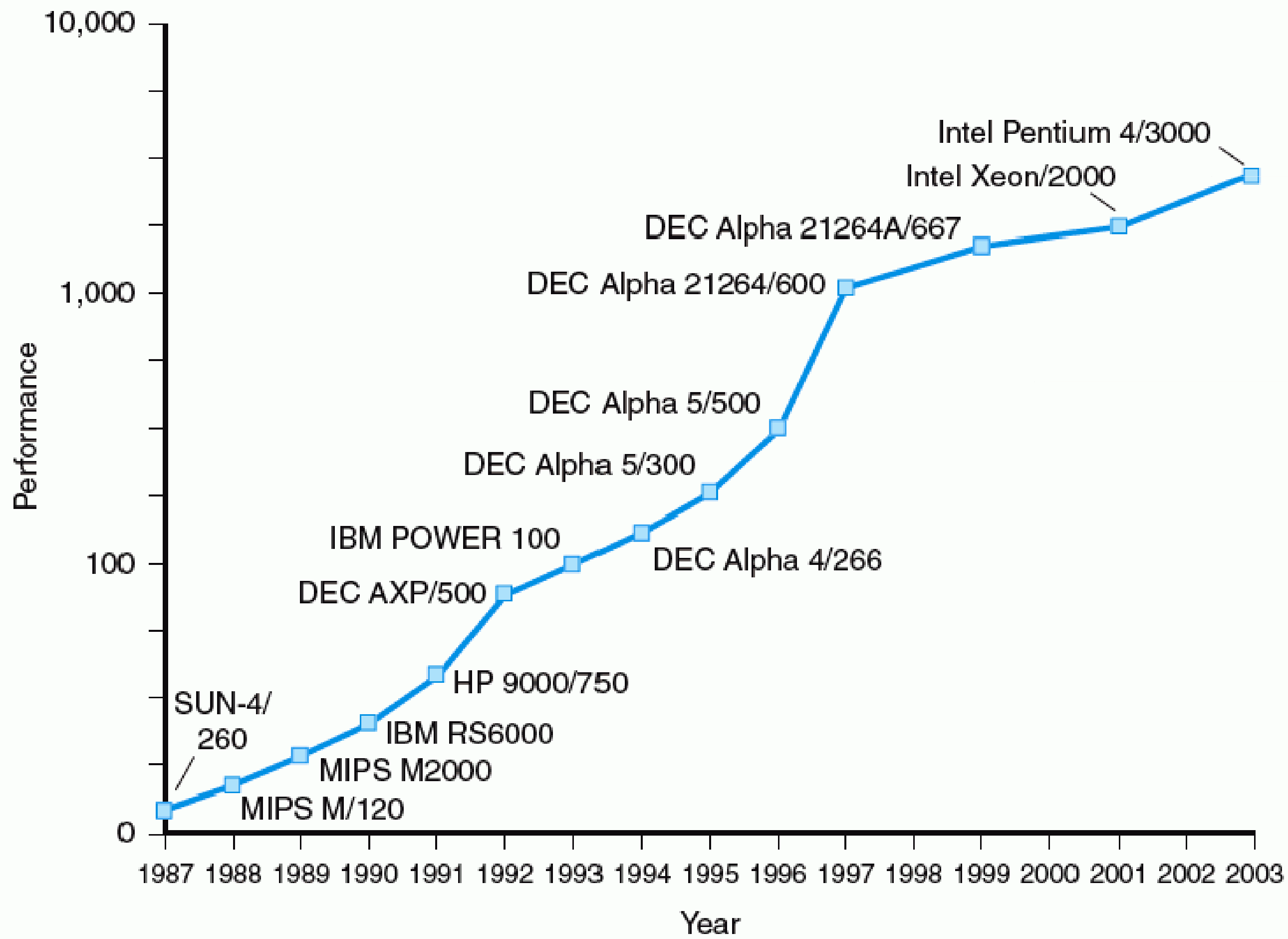


FIGURE 1.17 Performance increase of workstations, 1987–2003. Here performance is given as approximately the number of times faster than the VAX-11/780, which was a commonly used yardstick. The rate of performance improvement is between 1.5 and 1.6 times per year. These performance numbers are based on SPECint performance (see Chapter 2) and scaled over time to deal with changing benchmark sets. For processors listed with *x*/*y* after their name, *x* is the model number and *y* is the speed in megahertz.

Technology Trends

- Moore's Law
  - Continued (up until now, at least)

- **Absolute improvements in density, speed, power, costs**
  - SRAM: density: ~30% (annual), speed: ~20%
  - DRAM: density: ~60%, speed: ~4%
  - Disk: density: ~60%, speed: ~10% (non-transistor)
  - Big improvements in flash memory and network bandwidth, too

- **Changing quickly and with respect to each other!!**
  - Example: density increases faster than speed
  - Trade-offs are constantly changing
  - Re-evaluate/re-design for each technology generation

Technology Change Drives Everything

- Computers get **10x faster, smaller, cheaper** every 5-6 years!

- **Doubling every 1.5 years:**

  -memory capacity

  -processor speed (due to advances in technology <u>and</u> hardware organization)

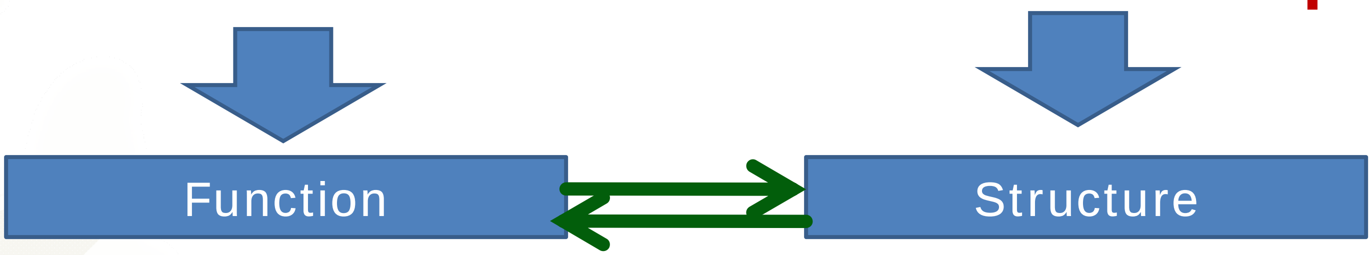
  - example: if Boeing had kept up with IBM we could *fly from Bangkok to Dhaka in 10 minutes for 500 Taka* !!

- New applications become self-sustaining
  - examples: mobile phones, digital cameras, mp3 players, etc.

Computer

- Computer is a complex system with millions of electronic components.

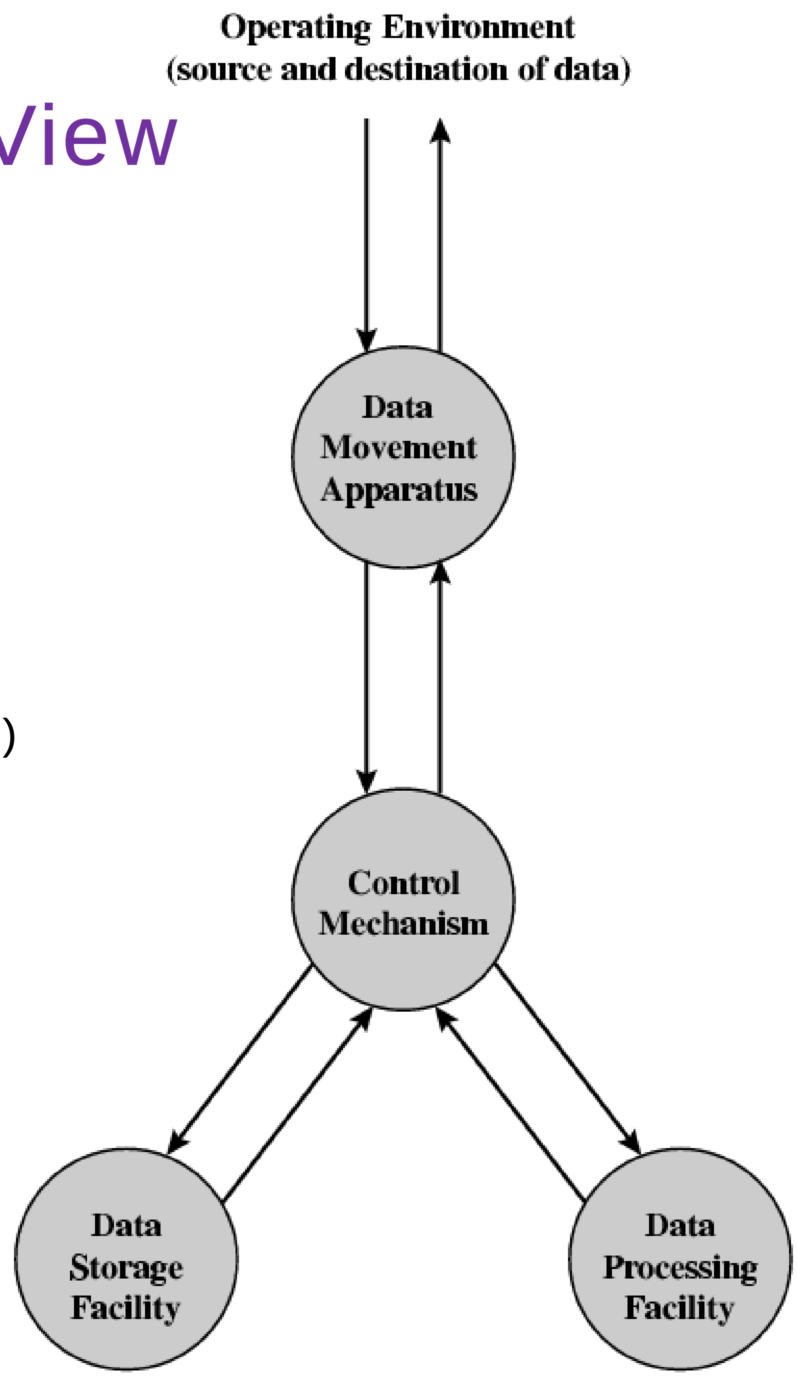
- Components are distributed into different hierarchical levels according to their **functionalities** and **interrelationship**.



Component's Operation

Way of interconnection between component's

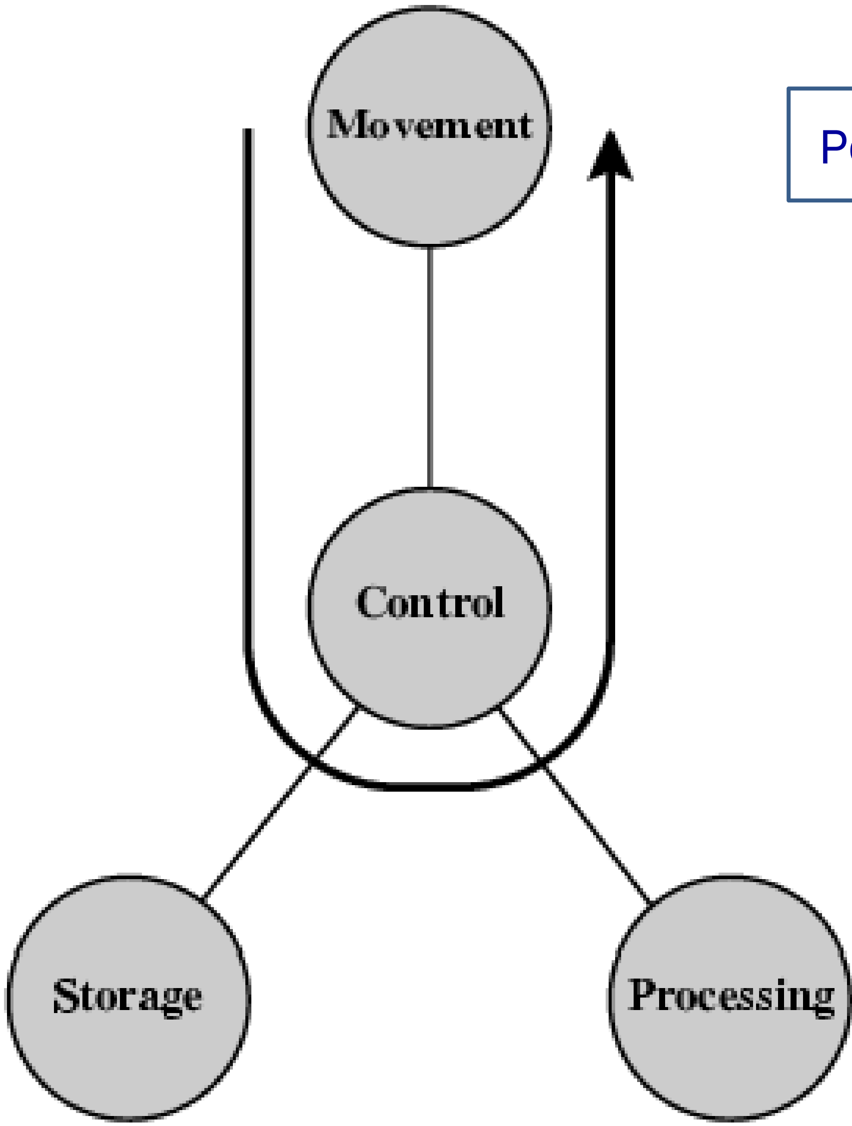
Computer : Functional View

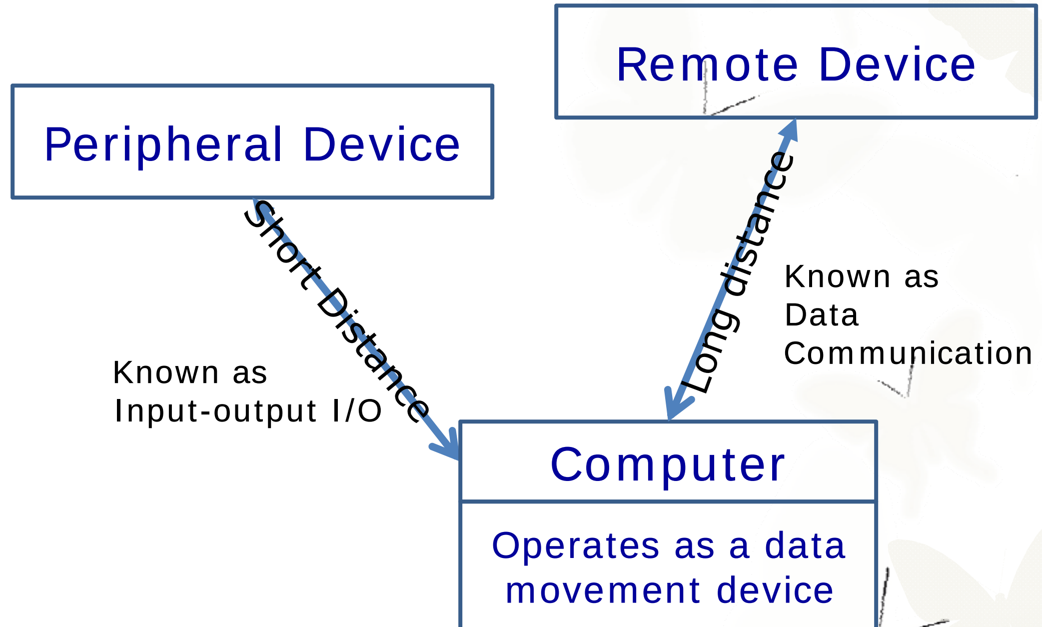


- Computer functions are categorized as
  - Data processing
  - Data storage (short term and long term)
  - Data movement
  - Control

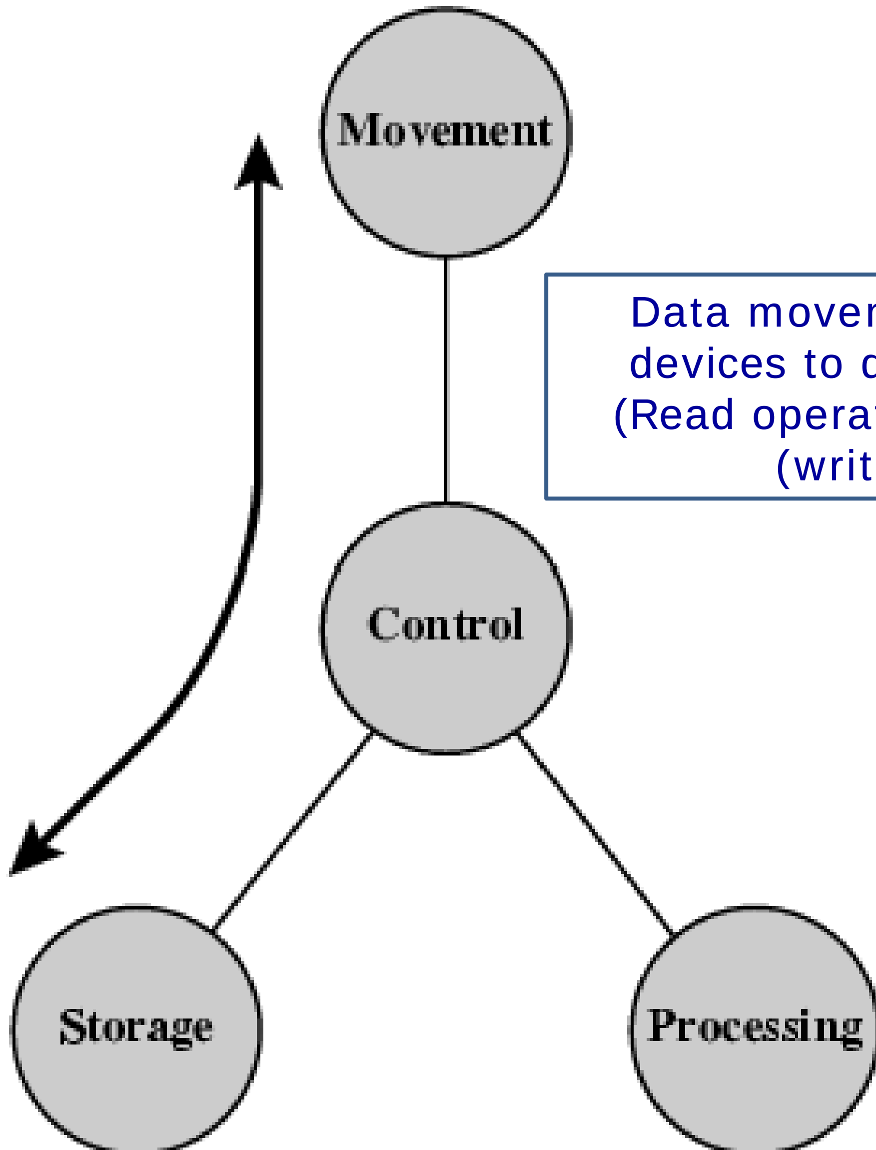
Four (4) possible operations are controlled by Control Mechanism

Operations Data movement
One device to another



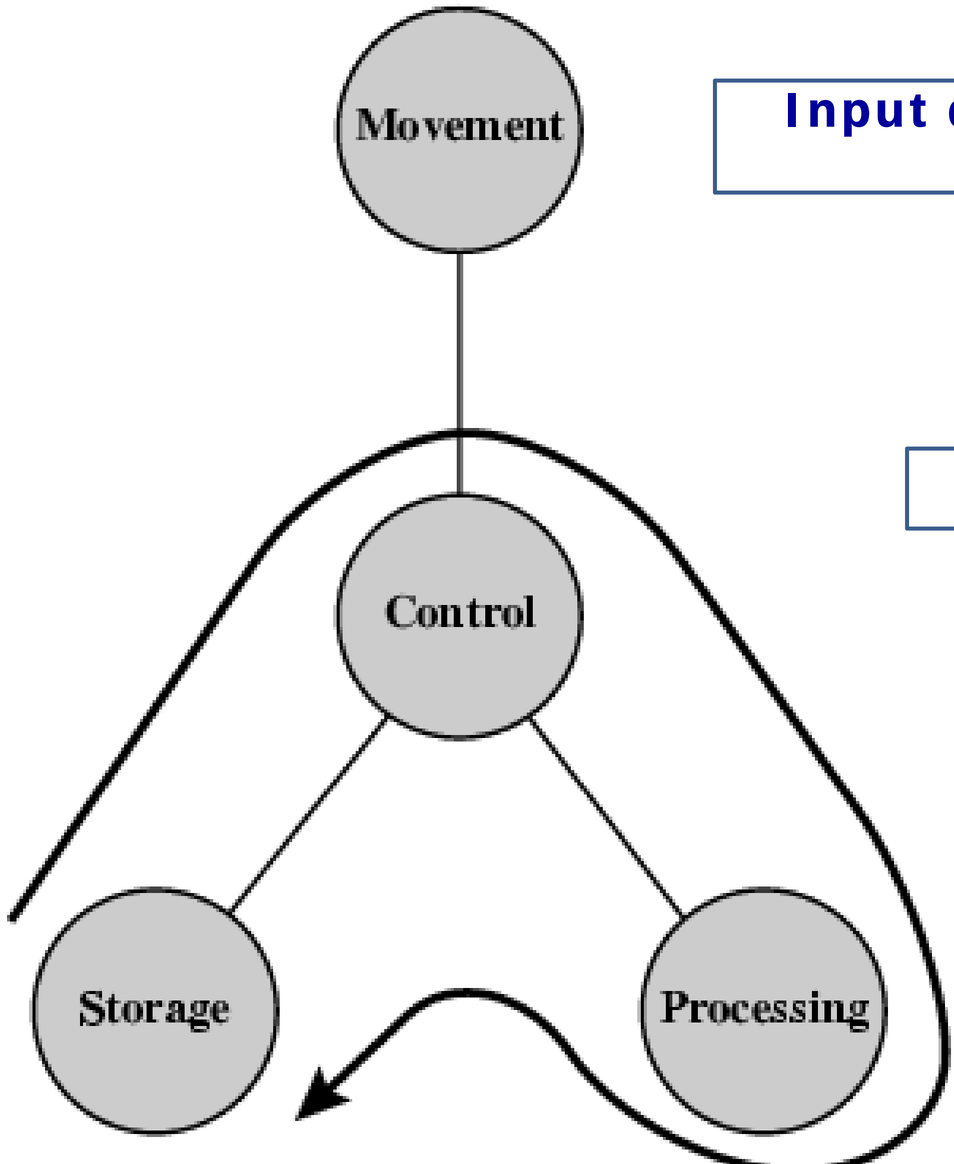


Operations Storage



Data movement from external devices to data storage device (Read operation) and vice versa (write operation)

Operation Processing from/to storage

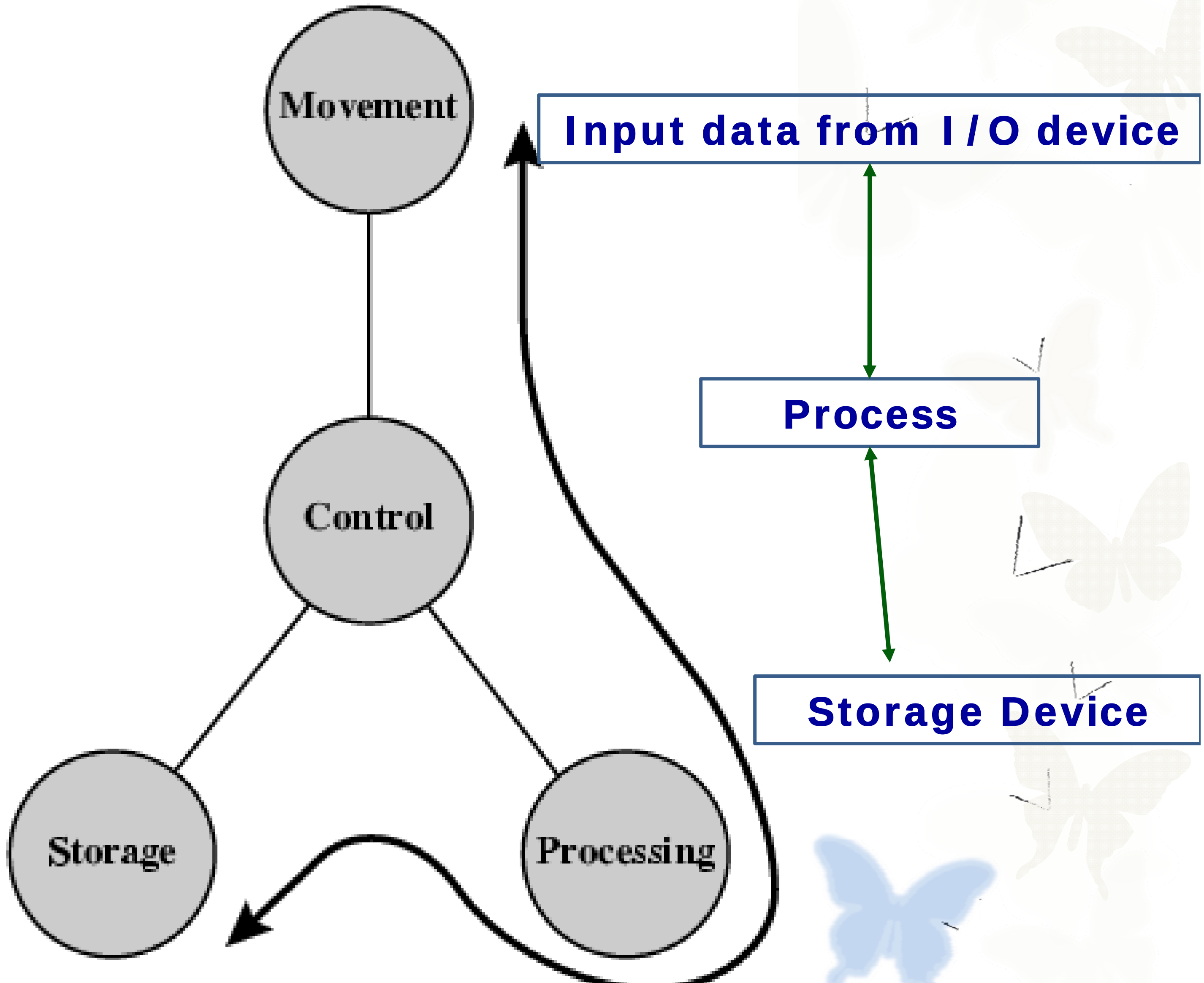


Input data from Storage Device

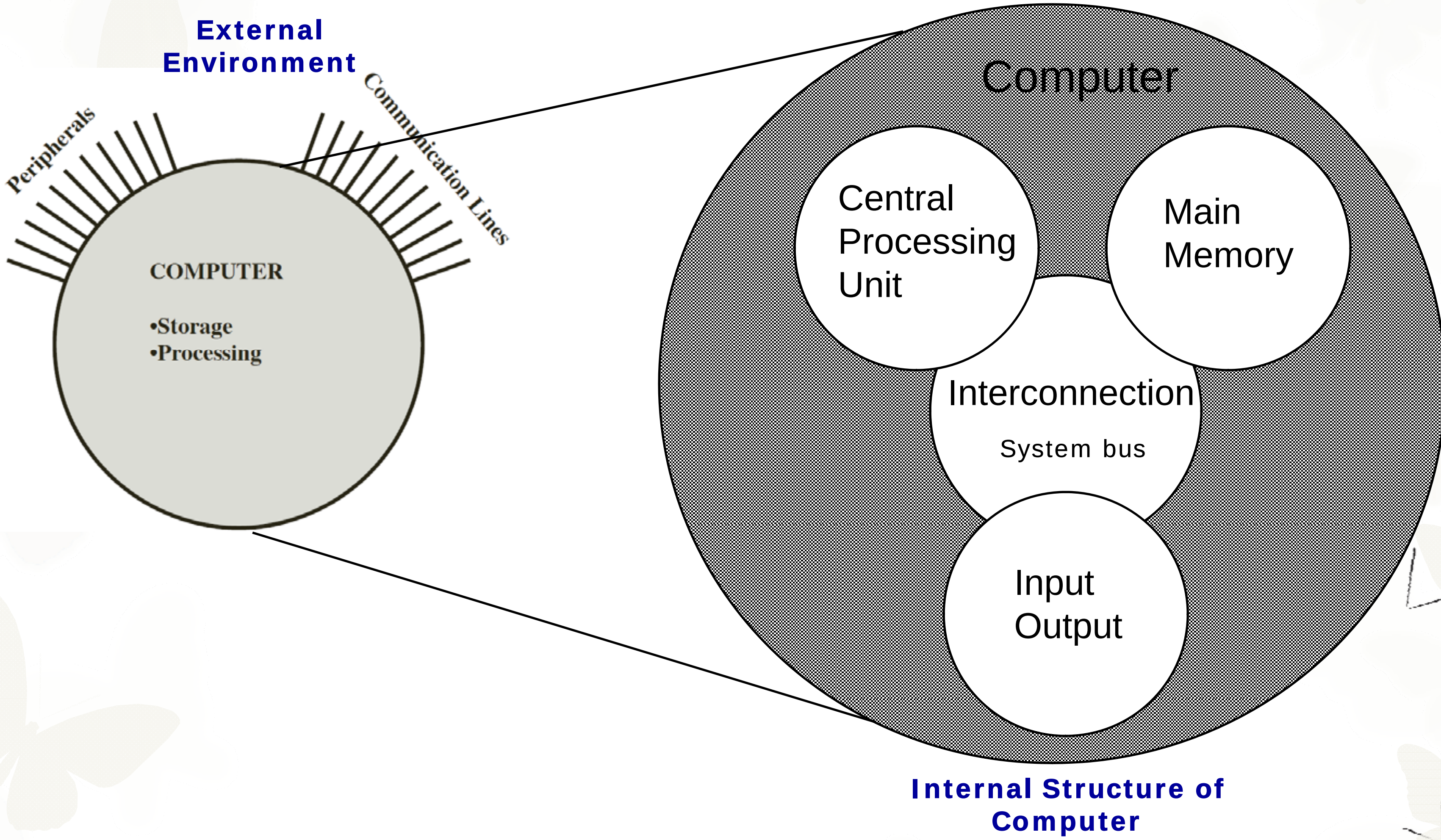
Process

Save data to Storage Device

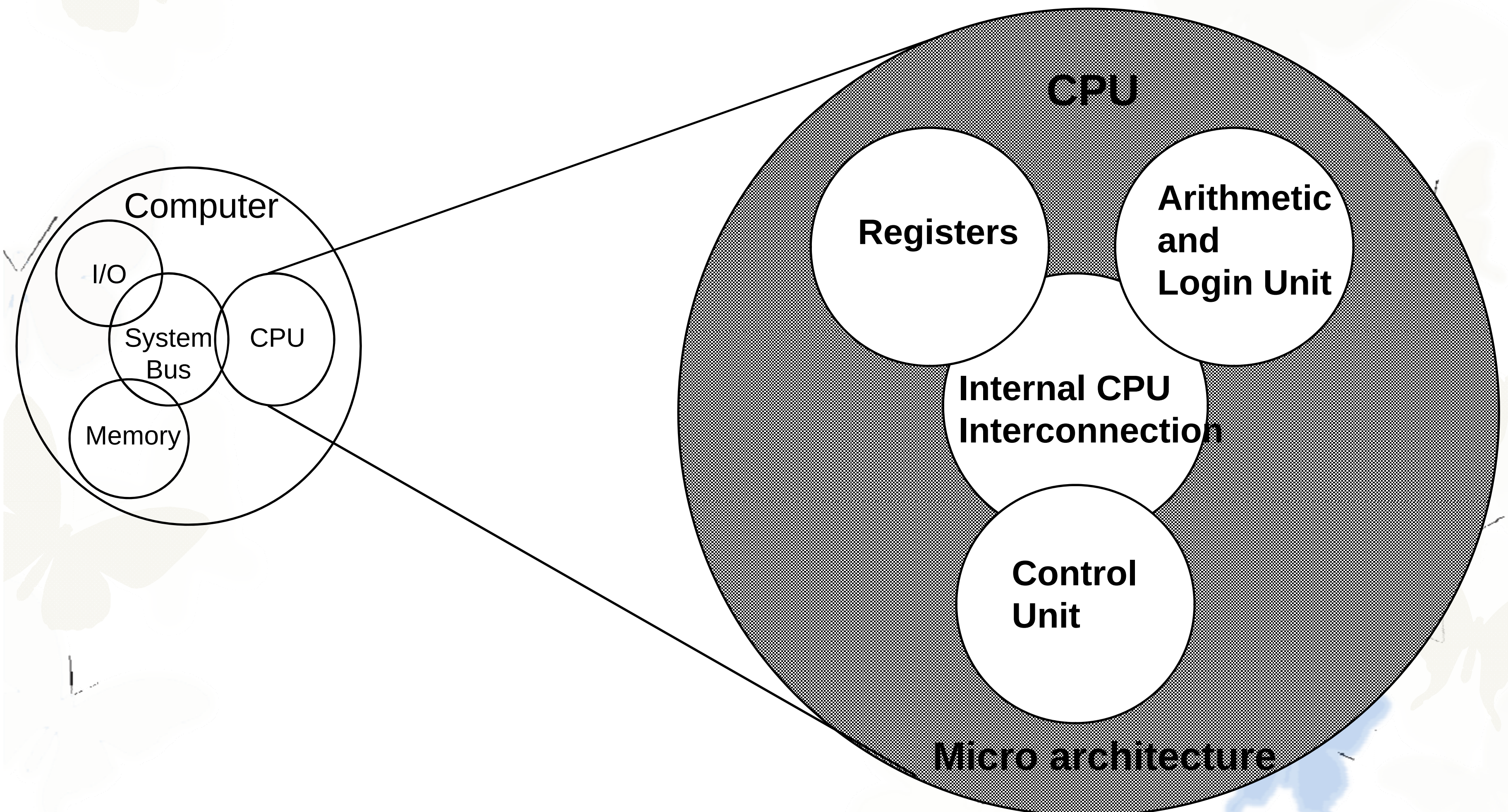
Operation Processing from storage to I/O



Computer and Top Level Structure

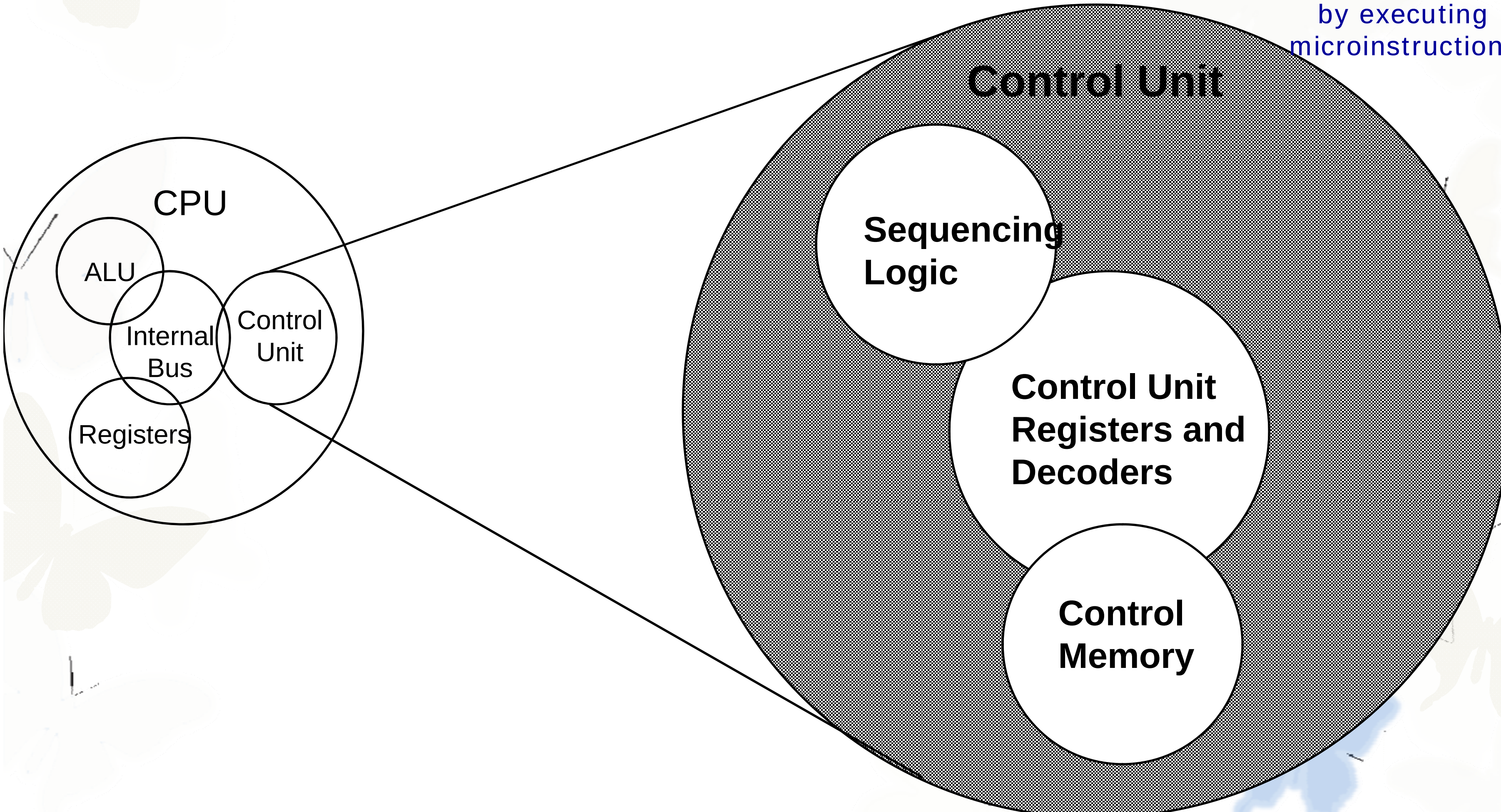


CPU Structure

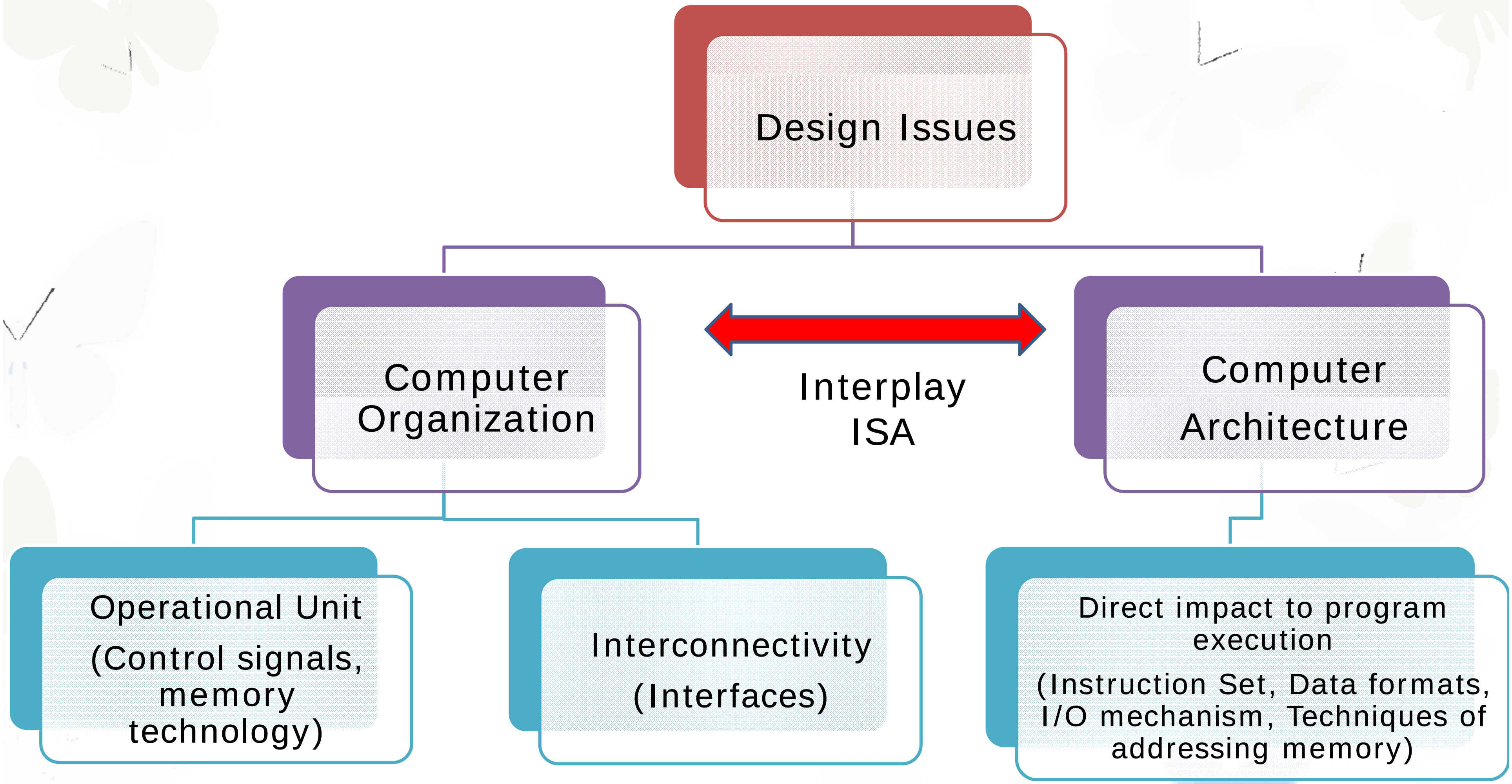


Control Unit Structure

Microprogrammed Implementation by executing microinstructions



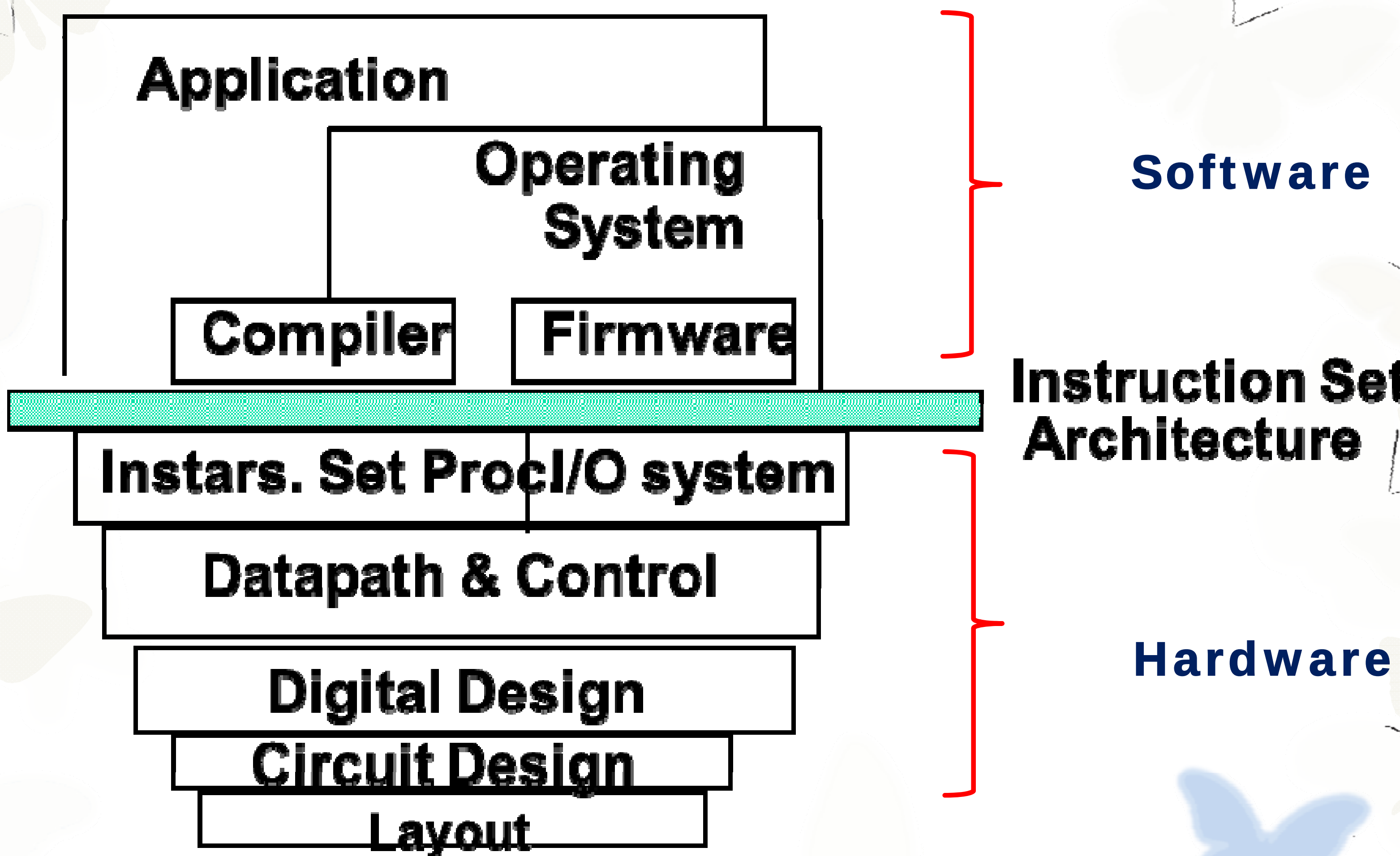
Computer Architecture Design



Computer Architecture Design: Example

- Architectural design issue:
  - whether a computer has a multiply instruction
- Organizational design issue:
  - whether the instruction execute by a special multiply unit or use repeated add unit of the system.
  - Based on frequency, relative speed, cost and physical size of the units.

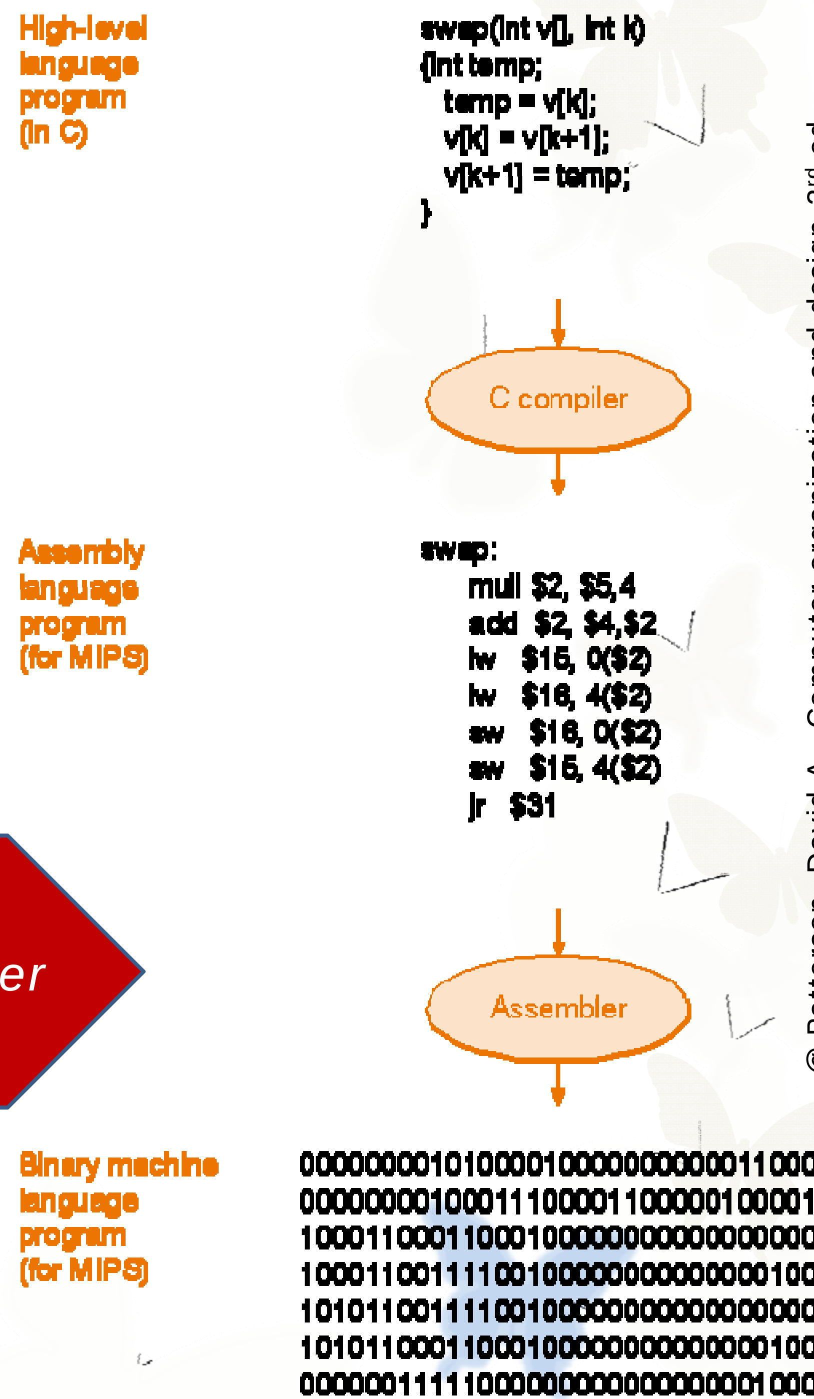
What is Computer Architecture (details)



Abstraction

- Investigating into the depths to reveal more information, but…
- An abstraction omits "unneeded" detail, helps us cope with complexity

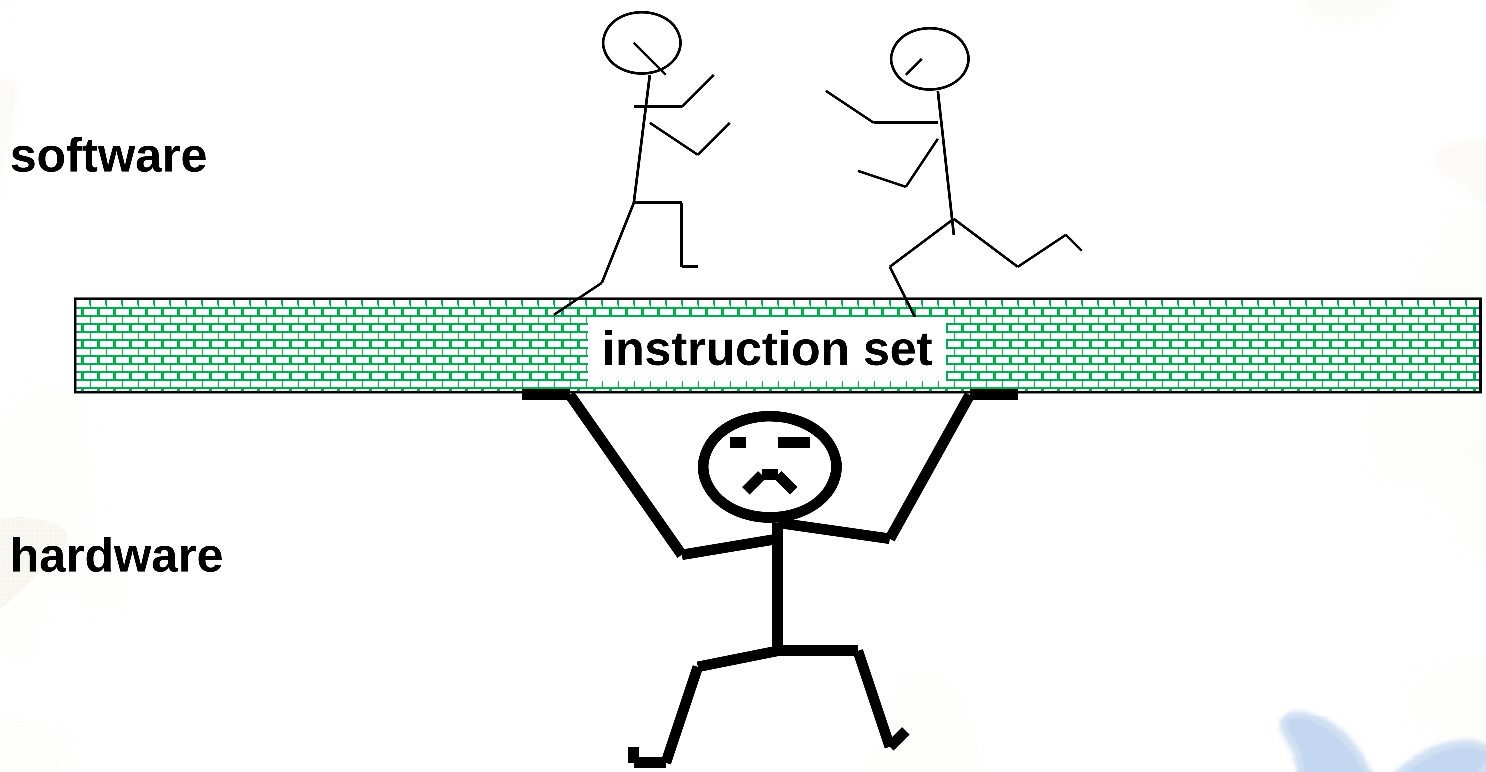
From the figure on the right,
How does abstraction help the programmer and how does she avoid too much detail?



@ Patterson, David A., Computer organization and design-3rd ed.

The Instruction Set: a Critical Interface

- is the agreed-upon interface between all the software that runs on the machine and the hardware that executes it.



The Instruction Set Architecture (ISA)

- that part of the architecture that is visible to the programmer
  - – operations-how many?, which one?
  - - operands –how many?, location
  - – number and types of registers
  - – instruction formats-size, formats
  - – storage access, addressing modes
  - – exceptional conditions

y=x+b

(add r1, r2, r5)

- advantage: *allows different implementations of the same architecture example: each instruction in MIPS is 32 bits*
- disadvantage: *sometimes prevents adding new innovations*

- Modern instruction set architectures:
  - - 80x86/Pentium, PowerPC, DEC Alpha, MIPS

Reading Assignment-1

- Read "*Cramming More Components onto Integrated Circuits*" by Gordon Moore

- **Answer the following question:**
  - One of the potential problems which Moore raises (and dismisses) is heat. Do you agree with Moore's conclusions? Either justify or refute Moore's conclusions.
  - A popular misconception of Moore's law is that it states that the speed of computers increases exponentially, however, that is not what Moore foretells in this paper.

  Explain what Moore's law actually says based on this paper.

Due Date: Wednesday 24/1/2018