# CSE 213
## Computer Architecture

# Lecture 1: History of Computer

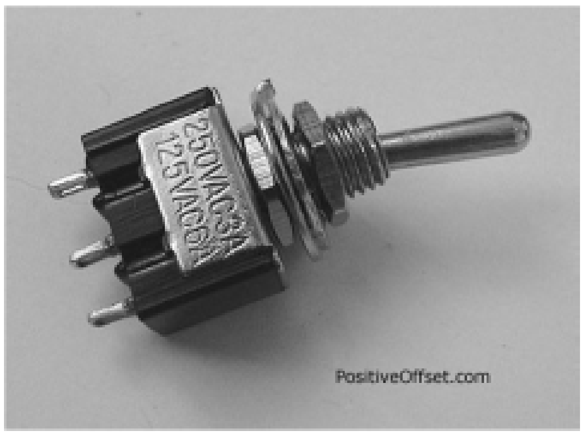
Military Institute of Science and Technology

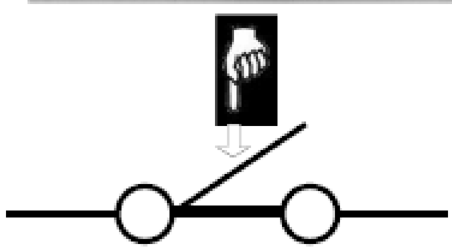
# Computer System Organization

The most amazing and likely to be most long-lived invention of the 1800’s was…

THE ELECTRIC SWITCH

# Basic Building Blocks: A switch



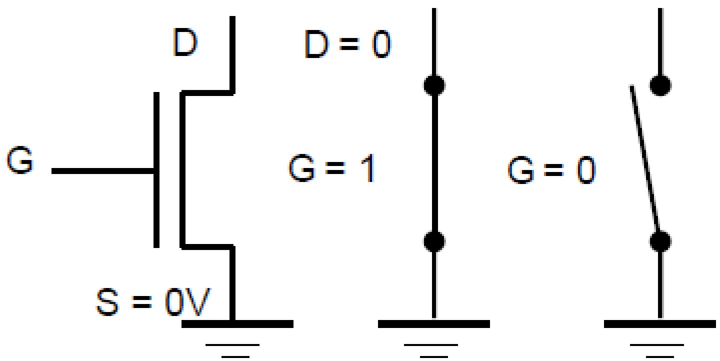



A switch is a simple device that can act as a conductor or isolator

Can be used for amazing things...

# NMOS and PMOS Transistors

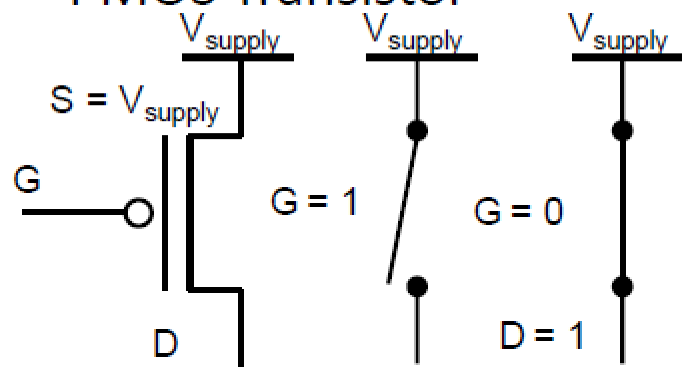
- NMOS Transistor




Closed switch
When $V_G = V_{supply}$

- Connect source to drain when gate = 1
- N-channel transistor

PMOS Transistor



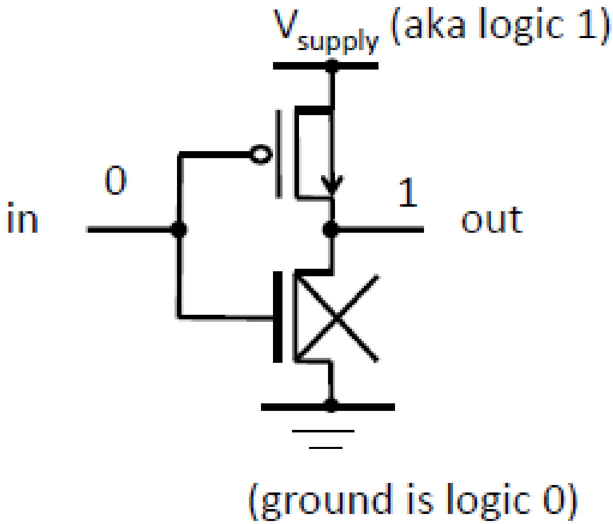

Closed switch
When $V_G$ = 0 V

Connect source to drain when gate = 0

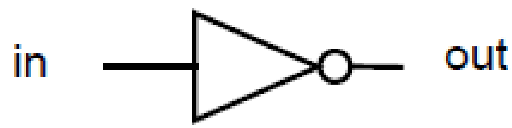
P-channel transistor

$V_S$: voltage at the source
$V_D$: voltage at the drain
$V_{supply}$: max voltage (aka a logical 1)
(ground): min voltage (aka a logical 0)

# Inverter




- Function: NOT
- Called an inverter
- Symbol:




| In | Out |
|---|---|
| 0 | 1 |
| 1 | 0 |

Truth table

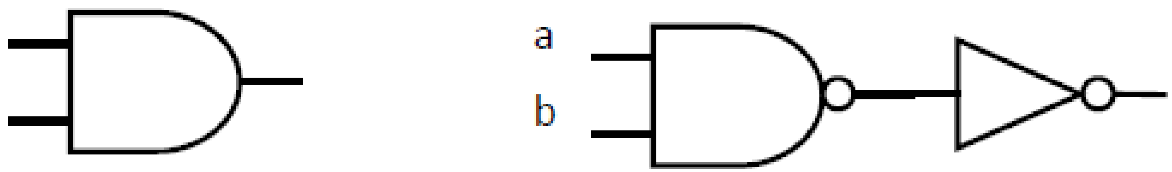
- Useful for taking the inverse of an input

- CMOS: complementary-symmetry metal–oxide–semiconductor

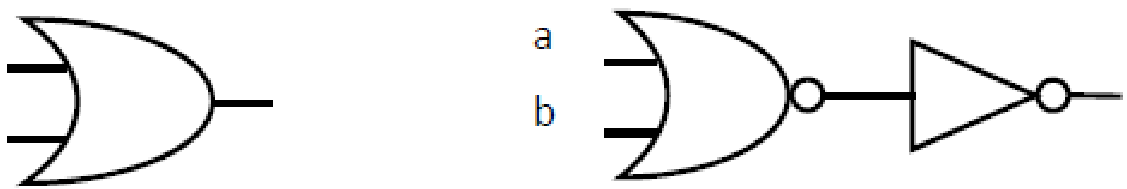
# Building Functions

NOT:

AND:



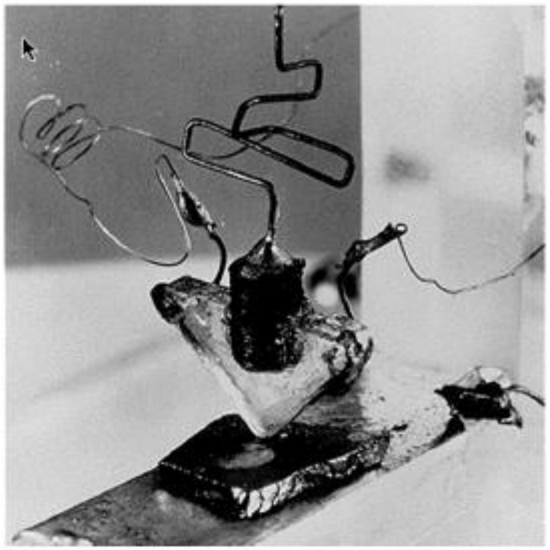

OR:

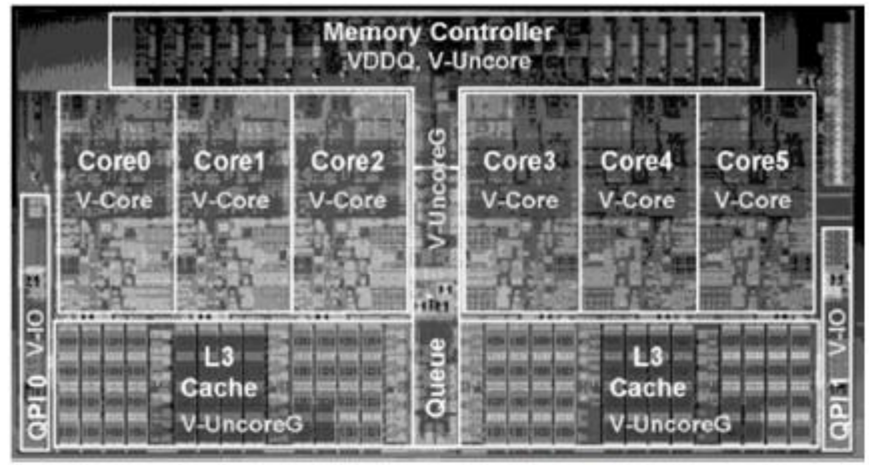



NAND and NOR are universal

- Can implement any function with NAND or just NOR gates
- useful for manufacturing

# Then and Now





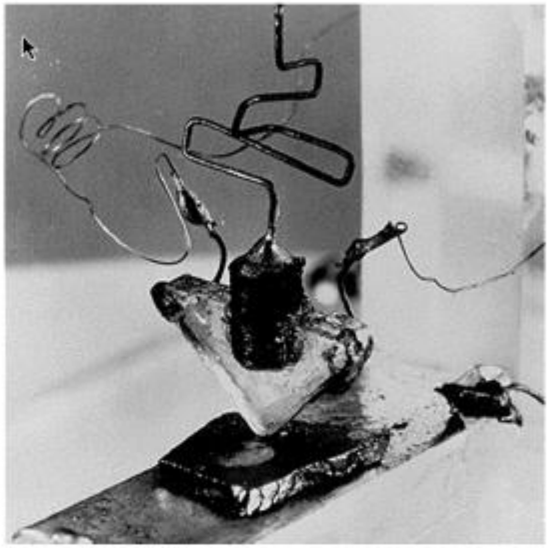

http://www.theregister.co.uk/2010/02/03/intel_westmere_ep_preview/

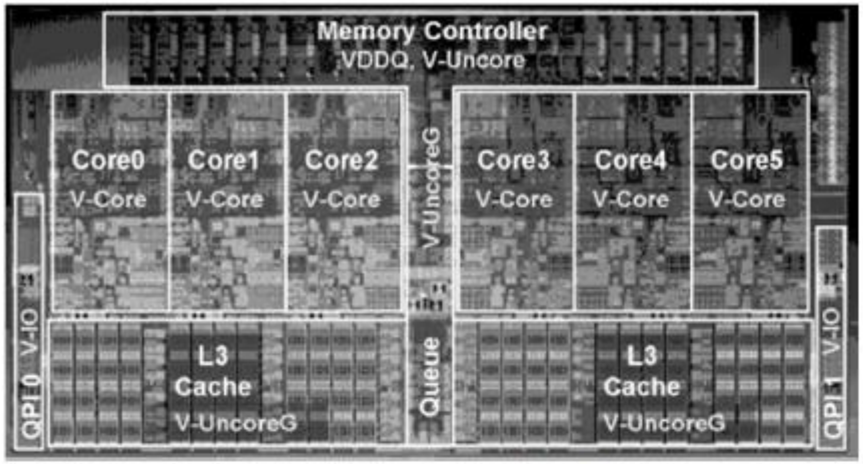
The first transistor

- on a workbench at
  AT&T Bell Labs in 1947
- Bardeen, Brattain, and Shockley

- An Intel Westmere
  - 1.17 billion transistors
  - 240 square millimeters
  - 32 nanometer: transistor gate width
  - Six processing cores
  - Release date: January 2010

# Then and Now





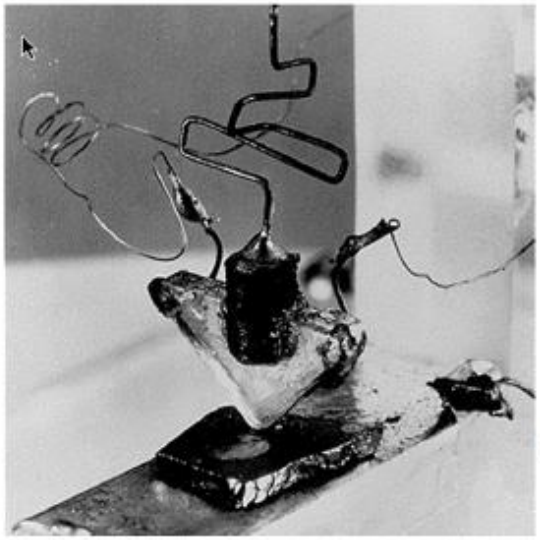

http://www.theregister.co.uk/2010/02/03/intel_westmere_ep_preview/

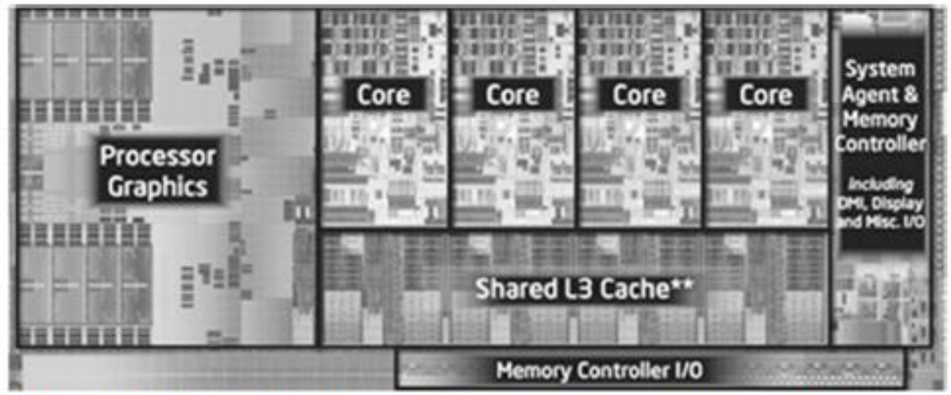
The first transistor

- on a workbench at
  AT&T Bell Labs in 1947
- Bardeen, Brattain, and Shockley

- An Intel Westmere
  - 1.17 billion transistors
  - 240 square millimeters
  - 32 nanometer: transistor gate width
  - Six processing cores
  - Release date: January 2010

# Then and Now





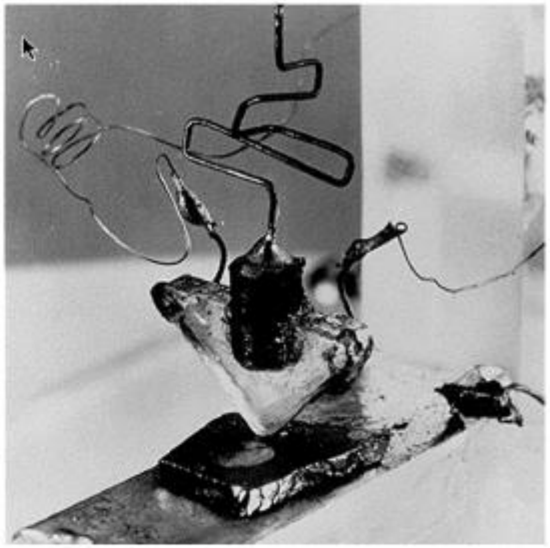

http://forwardthinking.pcmag.com/none/296972-intel-releases-ivy-bridge-first-processor-with-tri-gate-transistor

The first transistor

- on a workbench at
  AT&T Bell Labs in 1947
- Bardeen, Brattain, and Shockley

- An Intel Ivy Bridge
  - 1.4 billion transistors
  - 160 square millimeters
  - 22 nanometer: transistor gate width
  - Up to eight processing cores
  - Release date: April 2012

# Then and Now



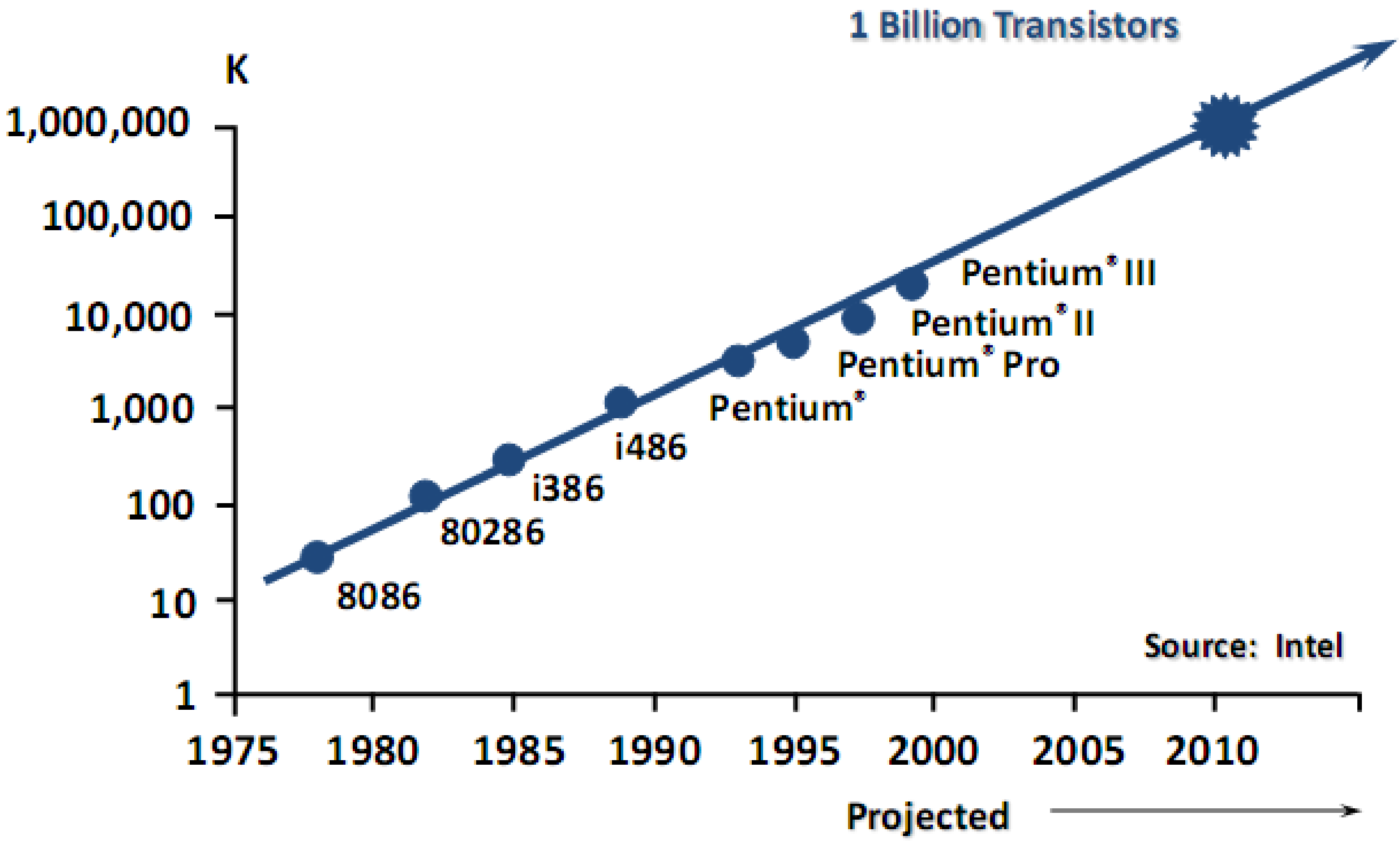

http://www.anandtech.com/show/6386/samsung-galaxy-note-2-review-t-mobile-/3

The first transistor

- on a workbench at
  AT&T Bell Labs in 1947
- Bardeen, Brattain, and Shockley

- Samsung Galaxy Note II
  - Eynos 4412 System on a Chip (SoC)
  - ARM Cortex-A9 processing core
  - 32 nanometer: transistor gate width
  - Four processing cores
  - Release date: November 2012

# Transistor Counts



# Moore's Law

The number of transistors integrated on a single die will double every 24 months...

– Gordon Moore, Intel co-founder, 1965

Amazingly Visionary

1971 – 2300 transistors – 1MHz – 4004

1990 – 1M transistors – 50MHz – i486

2001 – 42M transistors – 2GHz – Xeon

2004 – 55M transistors – 3GHz – P4

2007 – 290M transistors – 3GHz – Core 2 Duo

2009 – 731M transistors – 2GHz – Nehalem

2012 – 1400M transistors – 2-3GHz – Ivy Bridge

# Moore's Law

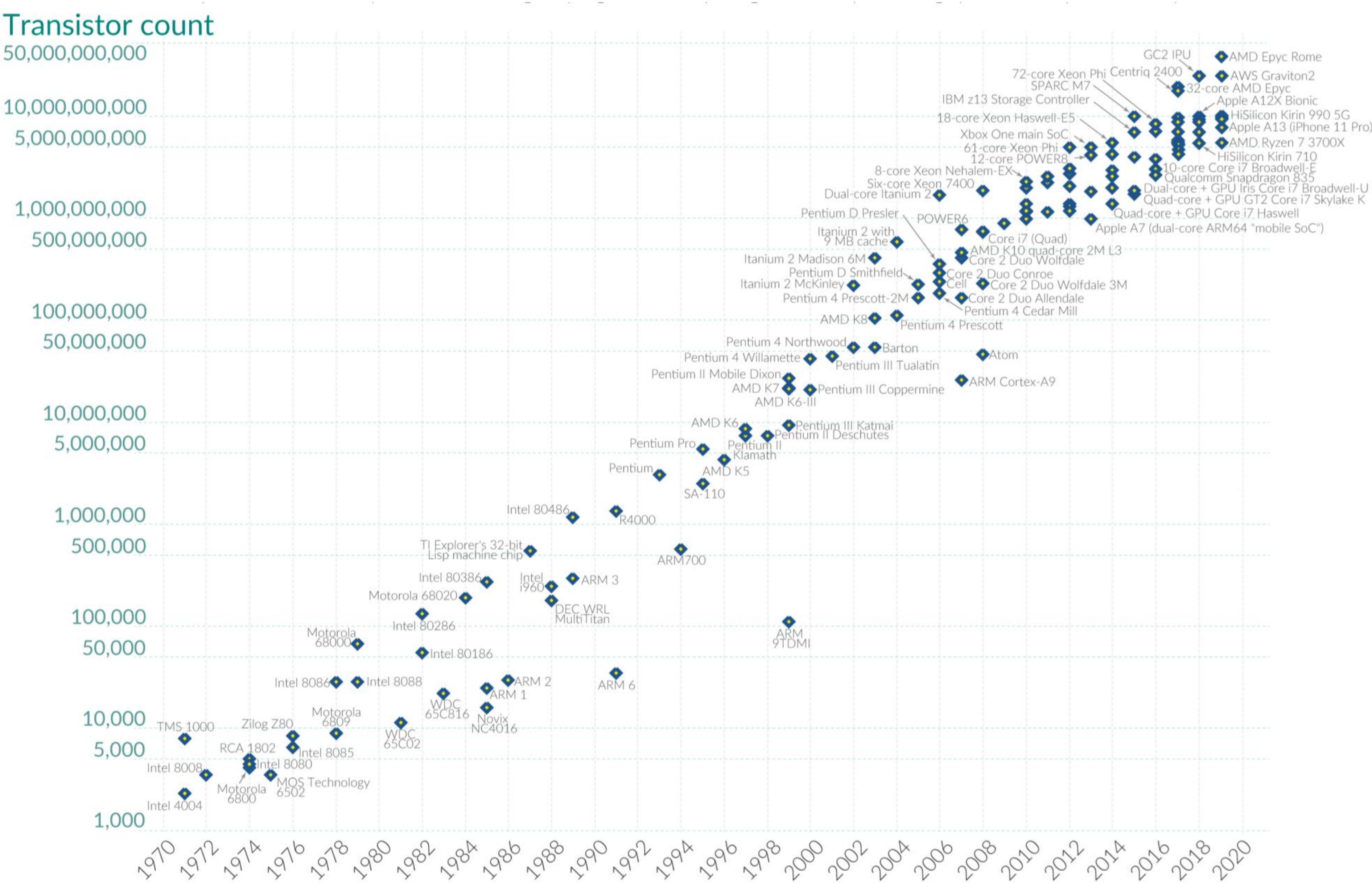
1965; Gordon Moore – co-founder of Intel

Observed number of transistors that could be put on a single chip was doubling every year

**The pace slowed to a doubling every 18 months in the 1970's but has sustained that rate ever since**

## Consequences of Moore's law:

- **The cost of computer logic and memory circuitry has fallen at a dramatic rate**
- **The electrical path length is shortened, increasing operating speed**
- **Computer becomes smaller and is more convenient to use in a variety of environments**
- **Reduction in power and cooling requirements**
- **Fewer interchip connections**



Transistor count
50,000,000,000
10,000,000,000
5,000,000,000
1,000,000,000
500,000,000
100,000,000
50,000,000
10,000,000
5,000,000
1,000,000
500,000
100,000
50,000
10,000
5,000
1,000
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
GC2 IPU
AMD Epyc Rome
72-core Xeon Phi
Centriq 2400
AWS Graviton2
SPARC M7
32-core AMD Epyc
IBM z13 Storage Controller
Apple A12X Bionic
18-core Xeon Haswell-E5
HiSilicon Kirin 990 5G
Apple A13 (iPhone 11 Pro)
Xbox One main SoC
61-core Xeon Phi
AMD Ryzen 7 3700X
12-core POWER8
HiSilicon Kirin 710
8-core Xeon Nehalem-EX
10-core Core i7 Broadwell-E
Six-core Xeon 7400
Qualcomm Snapdragon 835
Dual-core Itanium 2
Dual-core + GPU Iris Core i7 Broadwell-U
Quad-core + GPU GT2 Core i7 Skylake K
Pentium D Presler
POWER6
Quad-core + GPU Core i7 Haswell
Apple A7 (dual-core ARM64 "mobile SoC")
Itanium 2 with 9 MB cache
Core i7 (Quad)
Itanium 2 Madison 6M
AMD K10 quad-core 2M L3
Core 2 Duo Wolfdale
Pentium D Smithfield
Core 2 Duo Conroe
Itanium 2 McKinley
Cell
Core 2 Duo Wolfdale 3M
Pentium 4 Prescott-2M
Core 2 Duo Allendale
Pentium 4 Cedar Mill
AMD K8
Pentium 4 Prescott
Pentium 4 Northwood
Barton
Pentium 4 Willamette
Atom
Pentium III Tualatin
Pentium II Mobile Dixon
ARM Cortex-A9
AMD K7
Pentium III Coppermine
AMD K6-III
AMD K6
Pentium III Katmai
Pentium II Deschutes
Pentium Pro
Pentium II Klamath
Pentium
AMD K5
SA-110
Intel 80486
R4000
TI Explorer's 32-bit Lisp machine chip
ARM700
Intel 80386
Intel i960
ARM 3
Motorola 68020
DEC WRL MultiTitan
Intel 80286
ARM 9TDMI
Motorola 68000
Intel 80186
Intel 8086
Intel 8088
ARM 2
ARM 1
ARM 6
WDC 65C816
Novix NC4016
Motorola 6809
TMS 1000
Zilog Z80
WDC 65C02
RCA 1802
Intel 8085
Intel 8080
Intel 8008
MOS Technology 6502
Motorola 6800
Intel 4004

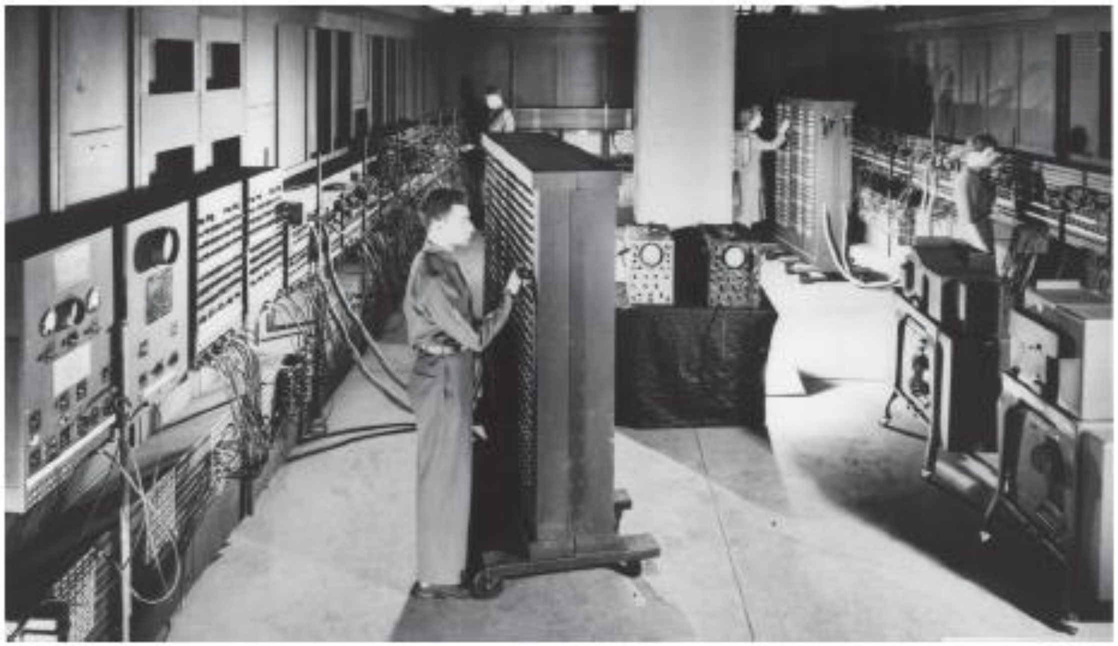
# Why large scale integration?

- Less die area, compactness
- More functions with less cost
- Less power consumption
- Less testing requirements at the system level
- Higher reliability, due to high quality on the chip interconnec
- Higher speed, due to reduced interconnect length
- Significant cost savings

# Computer Architecture

A bit of History

# The first electronic computers



ENIAC (Electronic Numerical Integrator and Calculator) –
The world's first electronic computer

3

# ENIAC

- ENIAC provided conditional jumps and was programmable, clearly distinguishing it from earlier calculators
- Programming was done manually by plugging cables and setting switches, and data was entered on punched cards. Programming for typical calculations required from half an hour to a whole day
- ENIAC was a general-purpose machine, limited primarily by a small amount of storage and tedious programming

# ENIAC

- J. Presper Eckert and John Mauchly at the Moore School of the University of Pennsylvania

- Funded by the United States Army

- Became operational during World War II but was not publicly disclosed until 1946

# ENIAC

Weighed 30 tons

Occupied 1500 square feet of floor space

Contained more than 18,000 vacuum tubes

140 kW Power consumption

Capable of 5000 additions per second

Decimal rather than binary machine

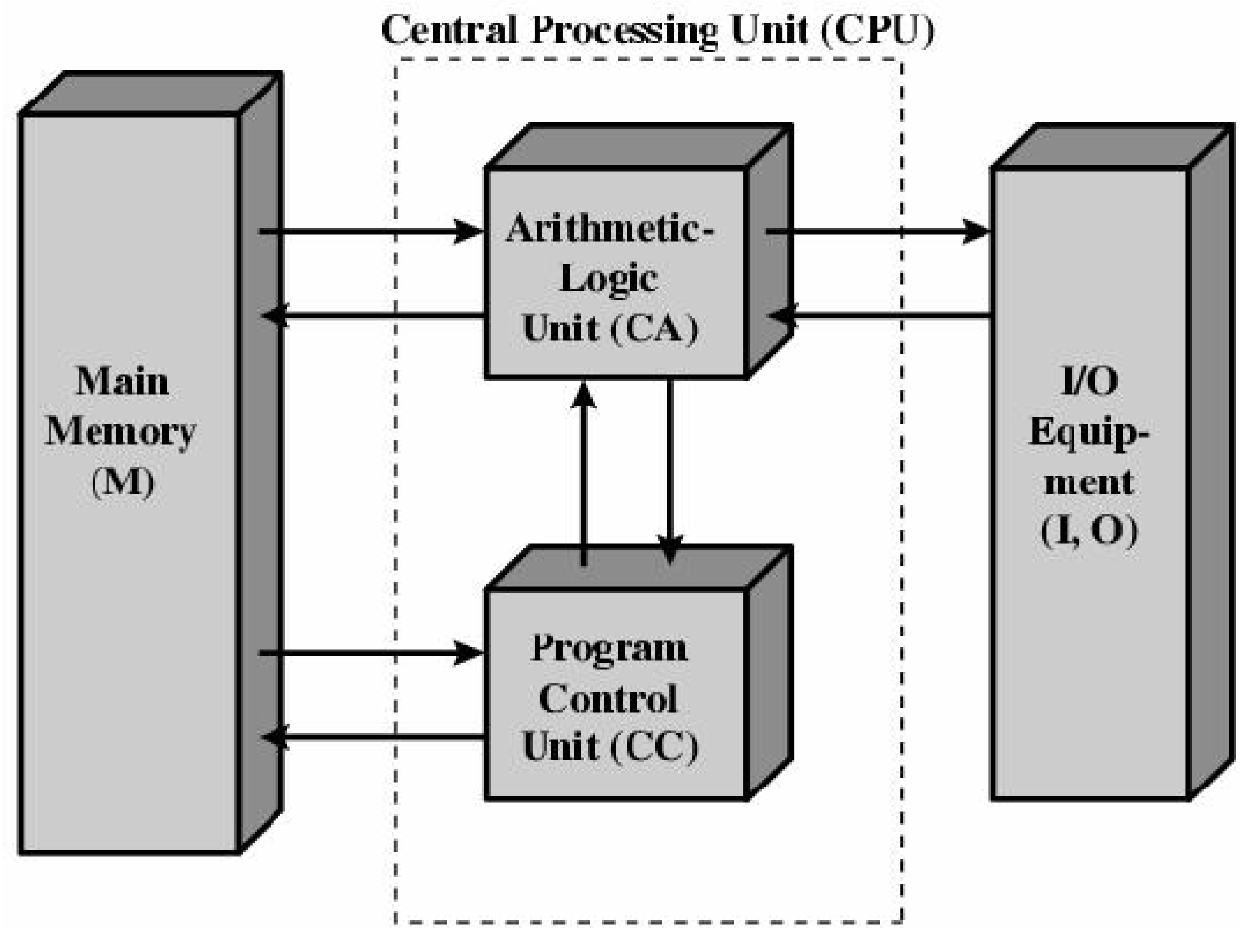
Memory consisted of 20 accumulators, each capable of holding a 10 digit number

Major drawback was the need for manual programming by setting switches and plugging/ unplugging cables

# The von Neumann Model

- The invention of stored program computers has been ascribed to a mathematician, John von Neumann, who was a contemporary of Mauchley and Eckert.

- Stored-program computers have become known as **von Neumann Architecture** systems.




**IAS (Princeton) computer model by Von Neumann's group.**

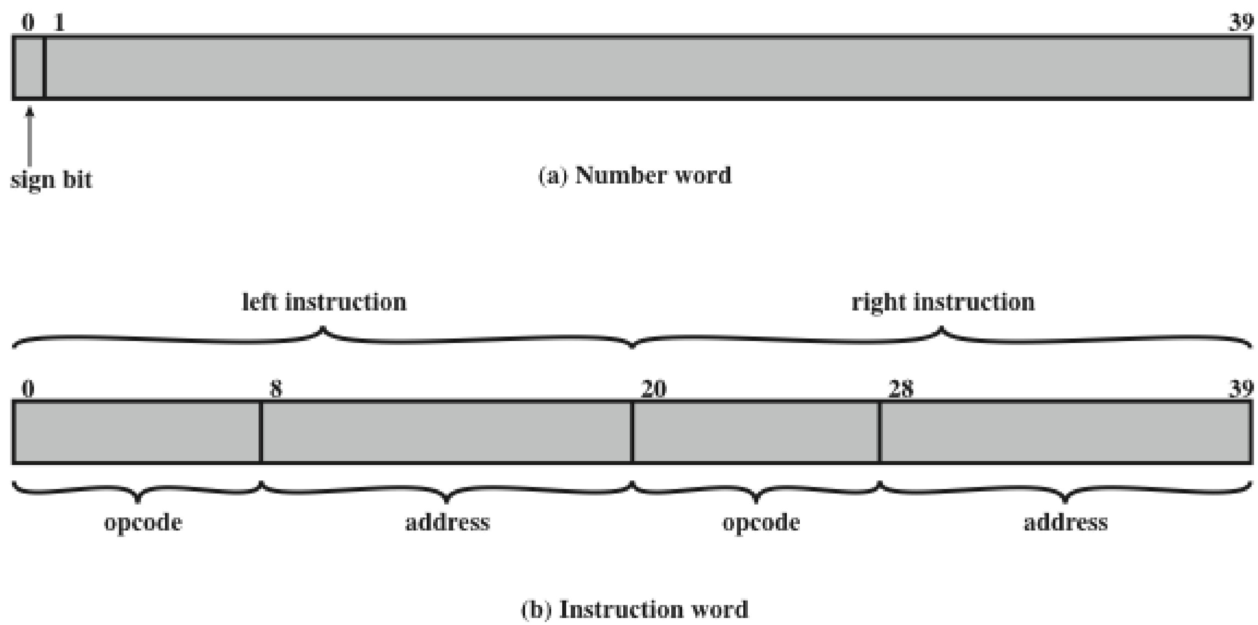
Institute for Advanced Studies(IAS)

+
# IAS Memory Formats

- The memory of the IAS consists of 1000 storage locations (called ***words***) of 40 bits each

- Both data and instructions are stored there

- Numbers are represented in binary form and each instruction is a binary code



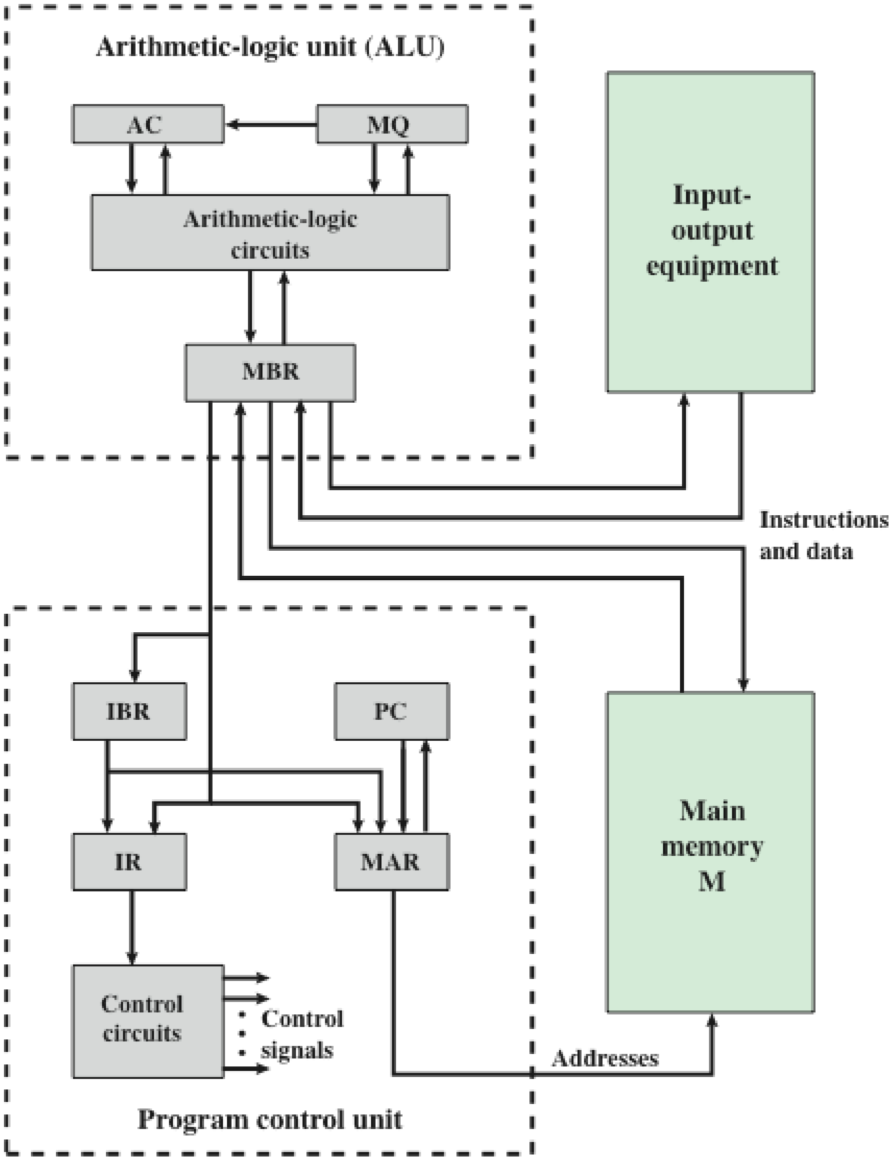
+

# Registers

| | |
|---|---|
| **Memory buffer register (MBR)** | • Contains a word to be stored in memory or sent to the I/O unit<br>• Or is used to receive a word from memory or from the I/O unit |
| **Memory address register (MAR)** | • Specifies the address in memory of the word to be written from or read into the MBR |
| **Instruction register (IR)** | • Contains the 8-bit opcode instruction being executed |
| **Instruction buffer register (IBR)** | • Employed to temporarily hold the right-hand instruction from a word in memory |
| **Program counter (PC)** | • Contains the address of the next instruction pair to be fetched from memory |
| **Accumulator (AC) and multiplier quotient (MQ)** | • Employed to temporarily hold operands and results of ALU operations |

+

# Structure of IAS Computer



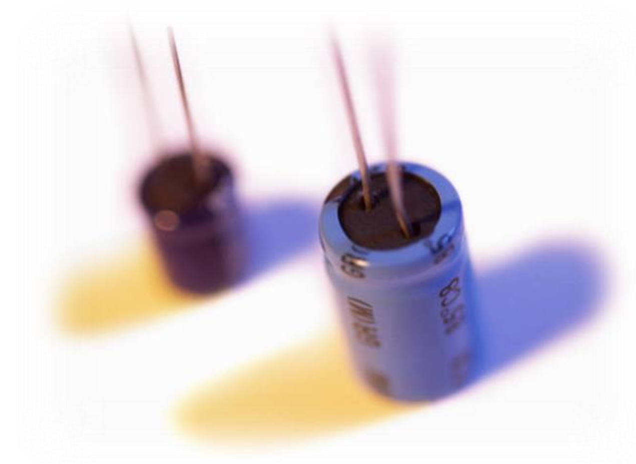
+

# Commercial Computers

## UNIVAC

- 1947 – Eckert and Mauchly formed the Eckert-Mauchly Computer Corporation to manufacture computers commercially
- UNIVAC I (Universal Automatic Computer)
  - First successful commercial computer
  - Was intended for both scientific and commercial applications
  - Commissioned by the US Bureau of Census for 1950 calculations
- The Eckert-Mauchly Computer Corporation became part of the UNIVAC division of the Sperry-Rand Corporation
- UNIVAC II – delivered in the late 1950's
  - Had greater memory capacity and higher performance
- Backward compatible

# History of Computers
## Second Generation: Transistors

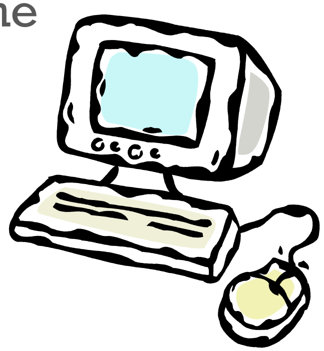


- Smaller
- Cheaper
- Dissipates less heat than a vacuum tube
- Is a *solid state device* made from silicon
- Was invented at Bell Labs in 1947
- It was not until the late 1950's that fully transistorized computers were commercially available

# Second Generation Computers

- Introduced:
  - More complex arithmetic and logic units and control units
  - The use of high-level programming languages
  - Provision of *system software* which provided the ability to:
    - load programs
    - move data to peripherals and libraries
    - perform common computations

- Appearance of the Digital Equipment Corporation (DEC) in 1957
- PDP-1 was DEC's first computer
- This began the mini-computer phenomenon that would become so prominent in the third generation



# History of Computers

## Third Generation: Integrated Circuits

- 1958 – the invention of the integrated circuit
- *Discrete component*
  - Single, self-contained transistor
  - Manufactured separately, packaged in their own containers, and soldered or wired together onto masonite-like circuit boards
  - Manufacturing process was expensive and cumbersome
- The two most important members of the third generation were the IBM System/360 and the DEC PDP-8

# Microelectronics

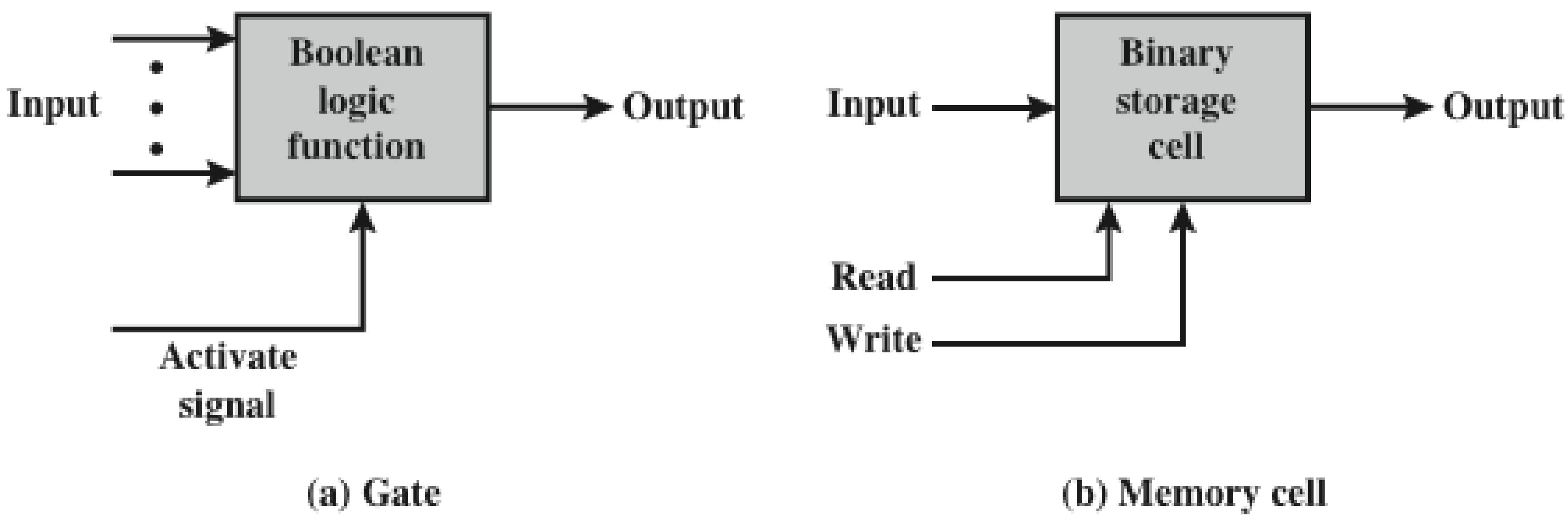



Figure 2.6 Fundamental Computer Elements

+

# Integrated Circuits

- Data storage – provided by memory cells
- Data processing – provided by gates
- Data movement – the paths among components are used to move data from memory to memory and from memory through gates to memory
- Control – the paths among components can carry control signals

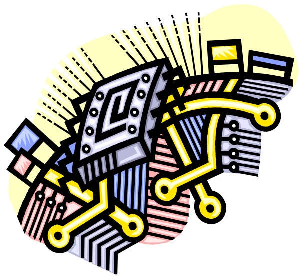
- A computer consists of gates, memory cells, and interconnections among these elements
- The gates and memory cells are constructed of simple digital electronic components
- Exploits the fact that such components as transistors, resistors, and conductors can be fabricated from a semiconductor such as silicon
- Many transistors can be produced at the same time on a single wafer of silicon
- Transistors can be connected with a processor metallization to form circuits

+

# Later Generations

**LSI**
Large
Scale
Integration

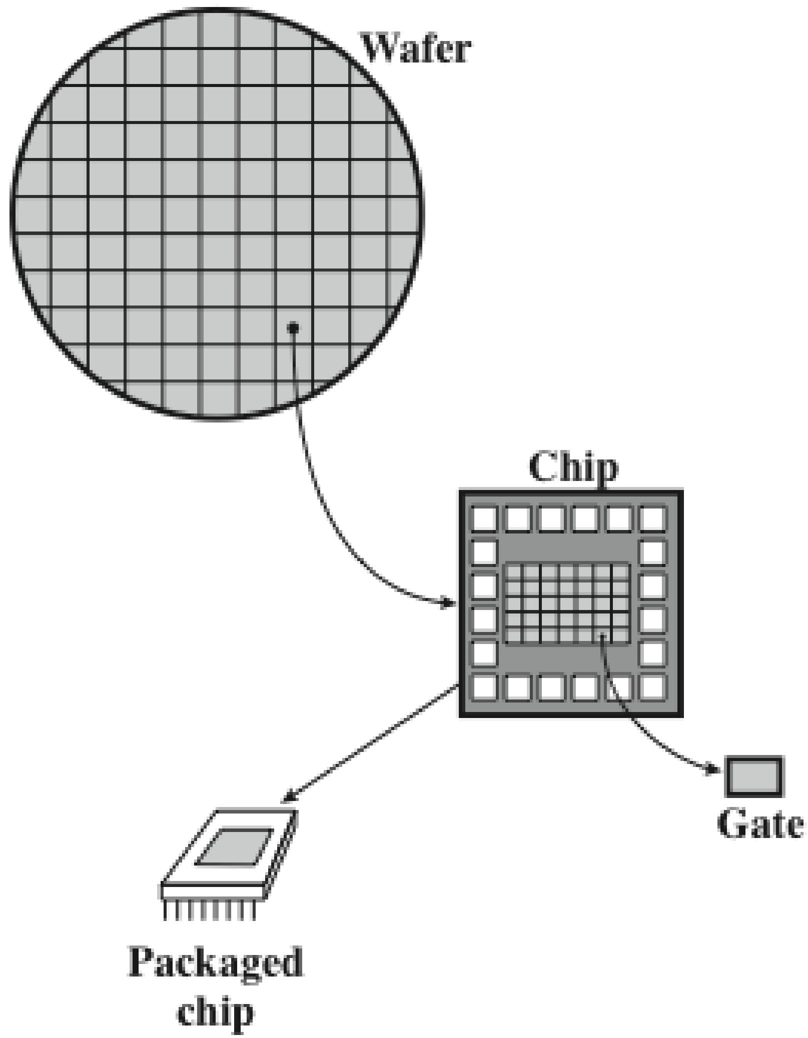
**VLSI**
Very Large
Scale
Integration



Semiconductor Memory
Microprocessors

**ULSI**
Ultra Large
Scale
Integration

# Wafer, Chip, and Gate Relationship



# Computer Generations

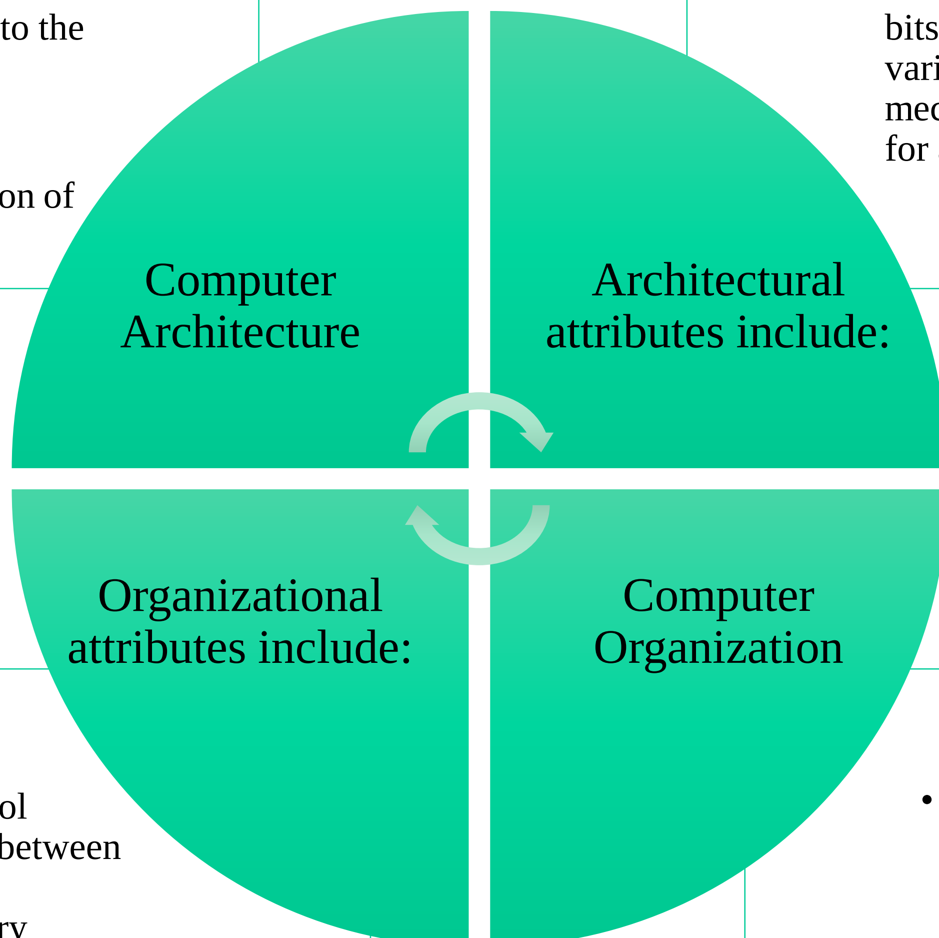
| Generation | Approximate Dates | Technology | Typical Speed (operations per second) |
|---|---|---|---|
| 1 | 1946–1957 | Vacuum tube | 40,000 |
| 2 | 1958–1964 | Transistor | 200,000 |
| 3 | 1965–1971 | Small and medium scale integration | 1,000,000 |
| 4 | 1972–1977 | Large scale integration | 10,000,000 |
| 5 | 1978–1991 | Very large scale integration | 100,000,000 |
| 6 | 1991- | Ultra large scale integration | 1,000,000,000 |

# Computer Architecture
# Vs
# Computer Organization

# Computer Architecture
# Computer Organization

- Attributes of a system visible to the programmer
- Have a direct impact on the logical execution of a program

- Instruction set, number of bits used to represent various data types, I/O mechanisms, techniques for addressing memory



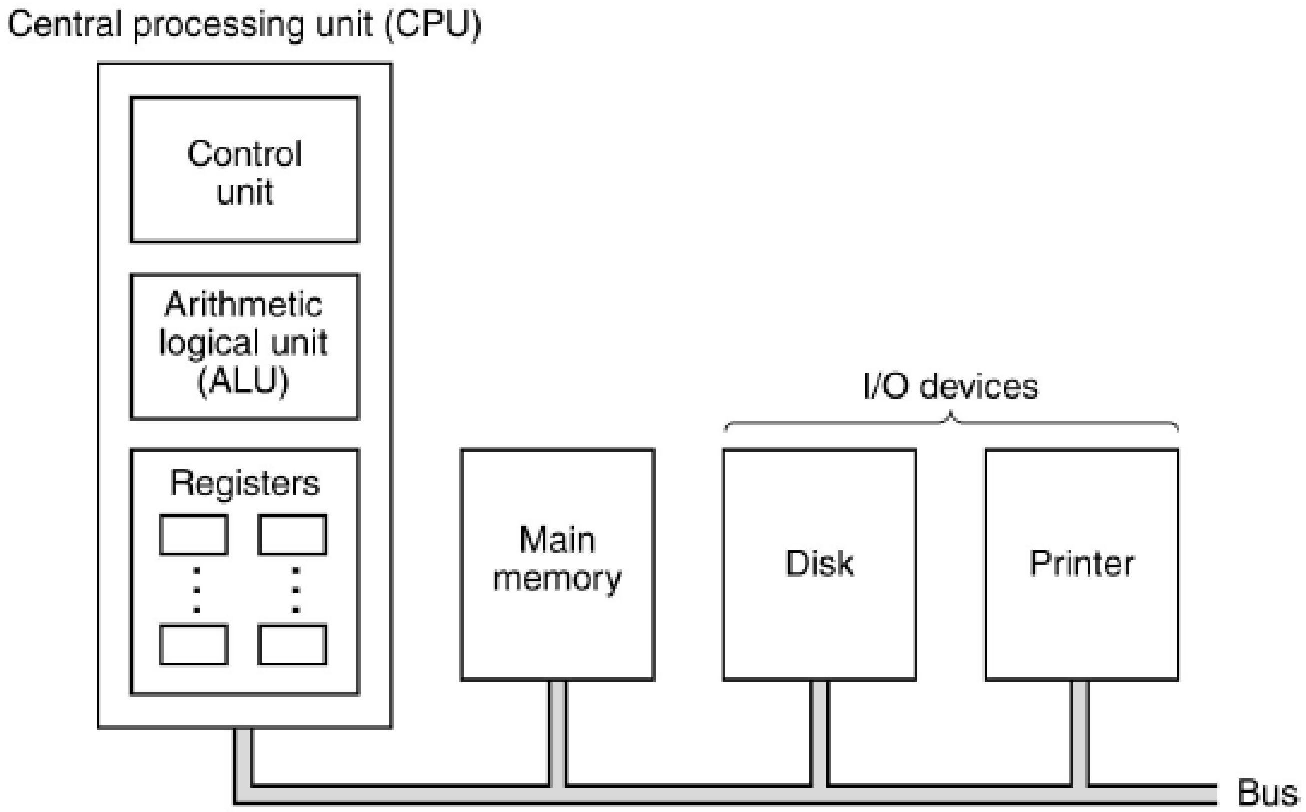

- Hardware details transparent to the programmer, control signals, interfaces between the computer and peripherals, memory technology used

- The operational units and their interconnections that realize the architectural specifications

# Computer Architecture

Basic Components

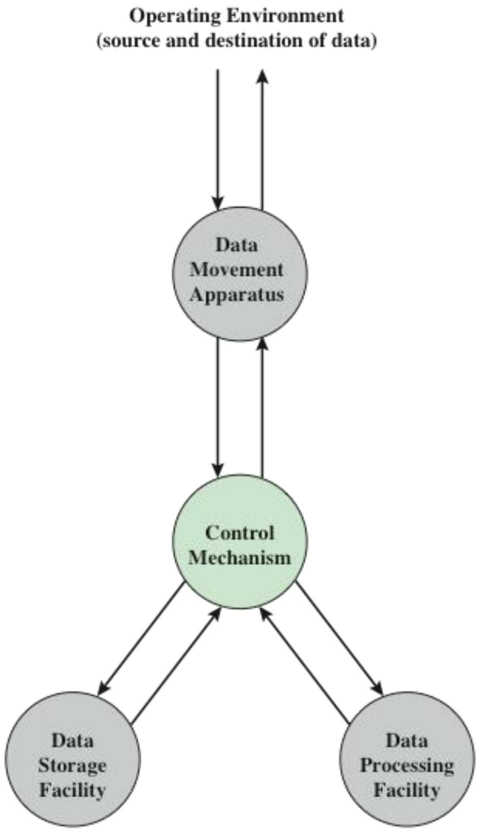
# Central Processing Unit (CPU) based CO




The organization of a simple computer with one CPU and two I/O devices

# Function

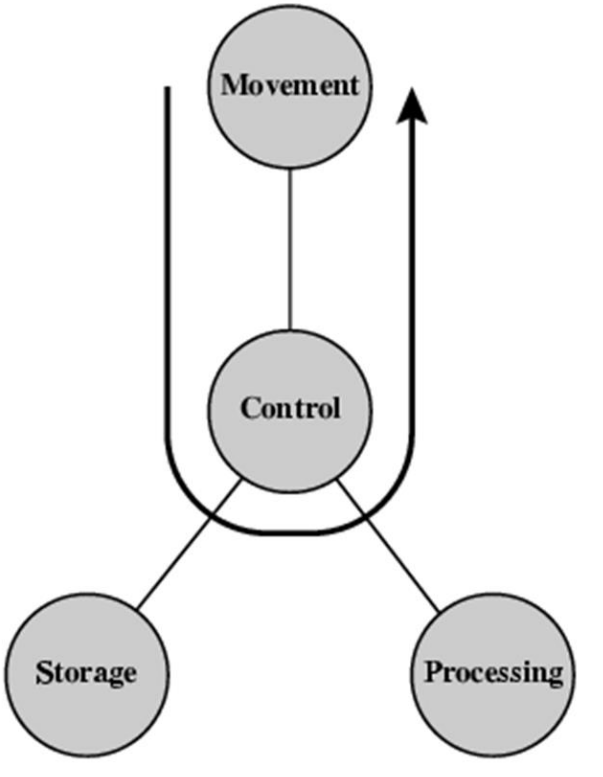
- A computer can perform four basic functions:
  a) Data movement
  b) Data storage
  c) Data processing
  d) Control



# Operation

**(a) Data movement:**

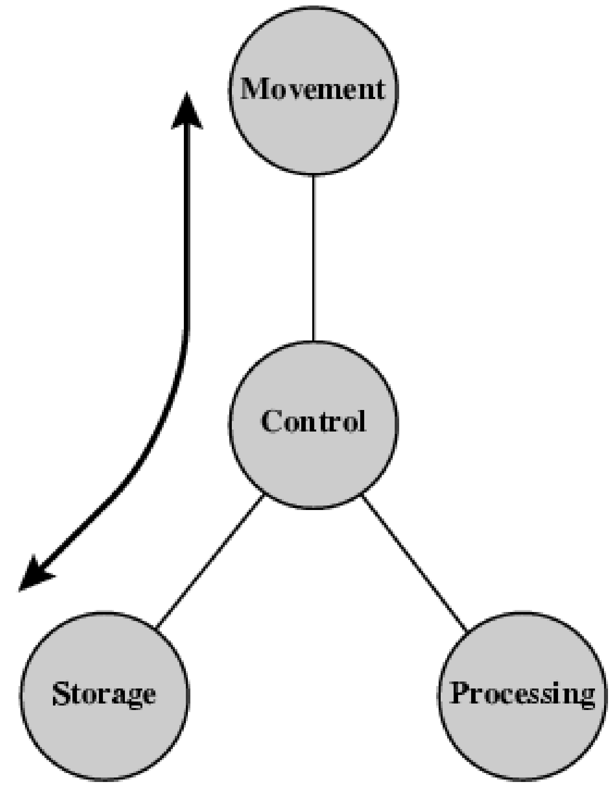
The computer must be able to **move data** between itself and the outside world. The computer's operating environment consists of devices that serve as either sources or destinations of data. When data are received from or delivered to a device that is directly connected to the computer, the process is known as *input–output (I/O), and the device is referred to as a peripheral*. *When data are moved* over longer distances, to or from a remote device, the process is known as *data communications*.



# Operation

**(b) Data Storage:**

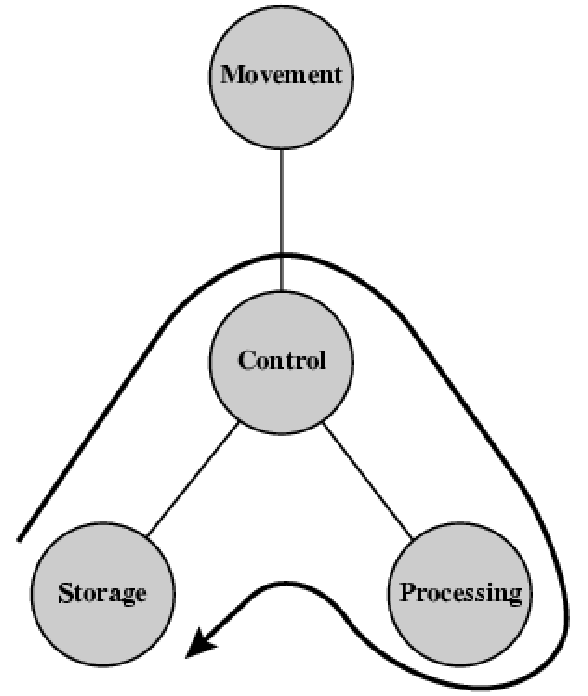
It is also essential that a computer **store data.** Even if the computer is processing data on the fly (i.e., data come in and get processed, and the results go out immediately), the computer must temporarily store at least those pieces of data that are being worked on at any given moment. Thus, there is at least a short-term data storage function. Equally important, the computer performs a long-term data storage function. Files of data are stored on the computer for subsequent retrieval and update.



# Operation

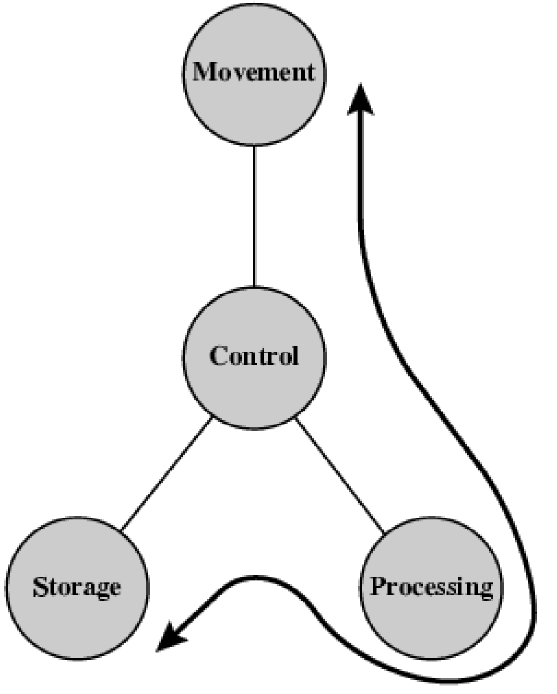
**(c) Data Processing:**

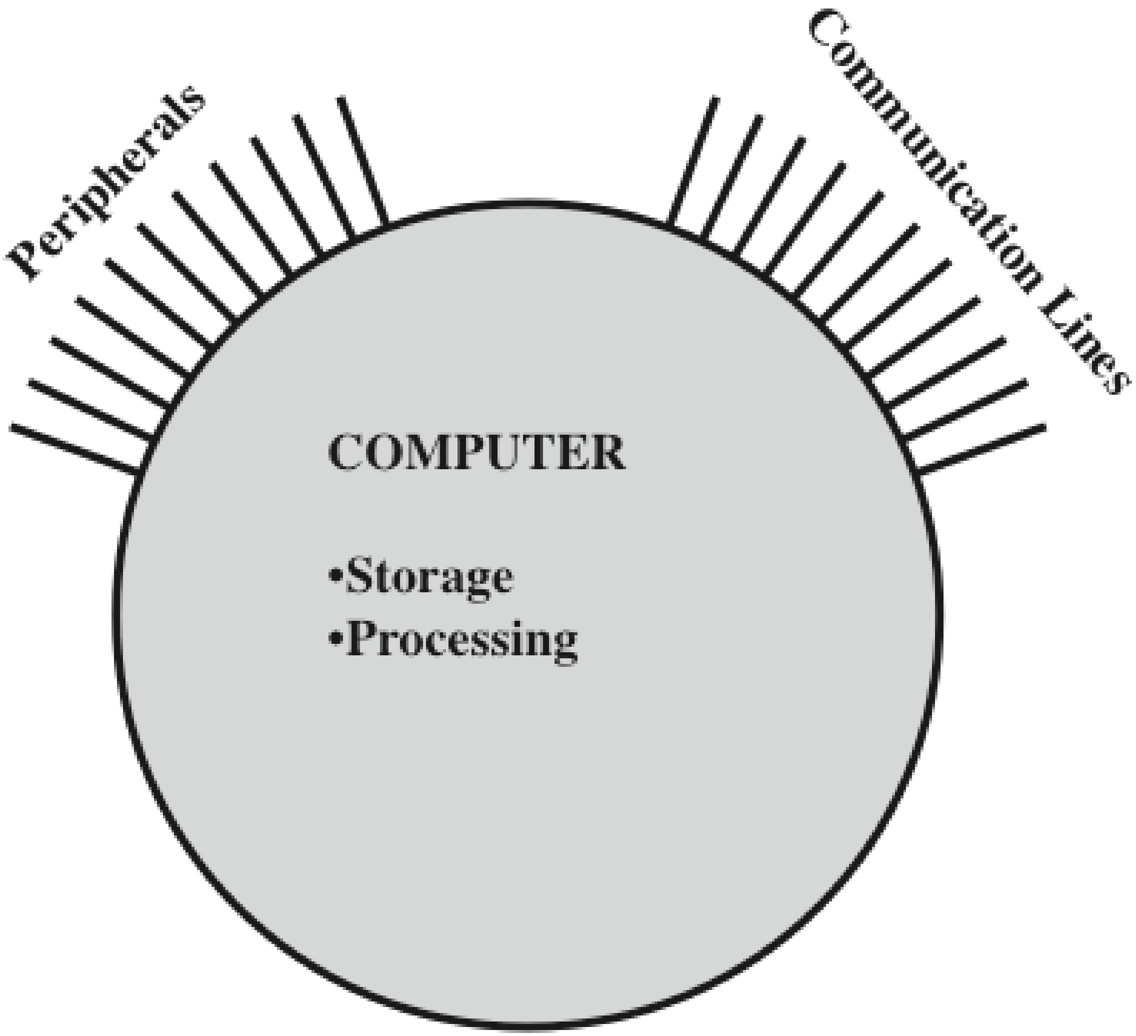
The computer, of course, must be able to **process data.** The data may take a wide variety of forms, and the range of processing requirements is broad. However, we shall see that there are only a few fundamental methods or types of data processing.



# Operation

**(d) Control:** There must be **control** of these three functions. Ultimately, this control is exercised by the individual(s) who provides the computer with instructions. Within the computer, a control unit manages the computer's resources and orchestrates the performance of its functional parts in response to those instructions.






# The Computer

# Structure and Function

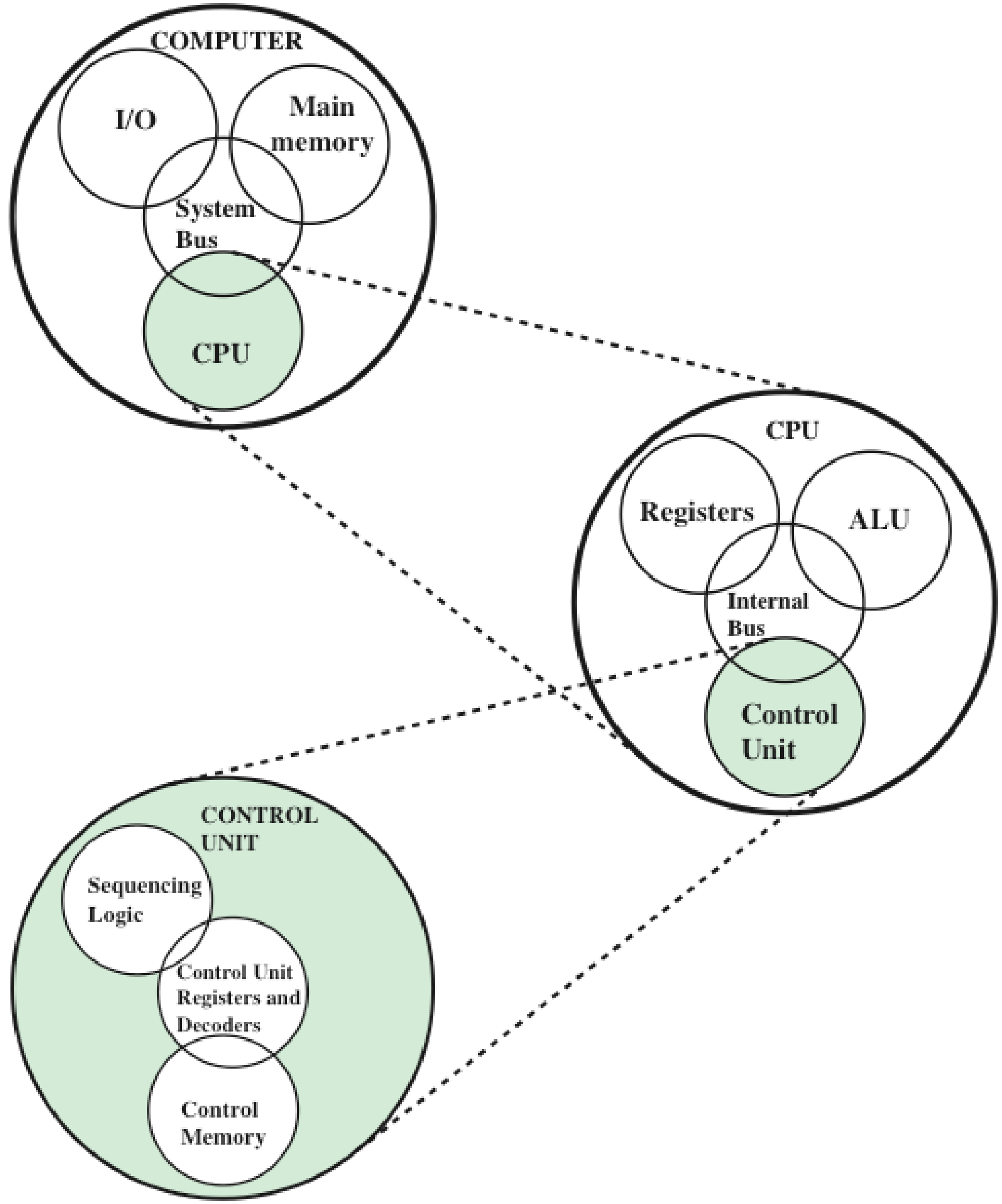
- Hierarchical system
  - Set of interrelated subsystems

– Hierarchical nature of complex systems is essential to both their design and their description

– Designer need only deal with a particular level of the system at a time
  - Concerned with structure and function at each level

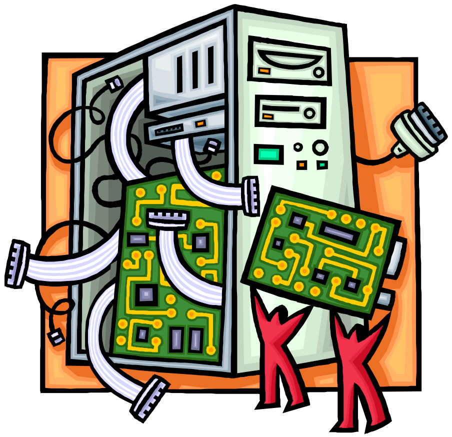
- Structure
  - The way in which components relate to each other
- Function
  - The operation of individual components as part of the structure

# Structure



COMPUTER
I/O
Main memory
System Bus
CPU
CPU
Registers
ALU
Internal Bus
Control Unit
CONTROL UNIT
Sequencing Logic
Control Unit Registers and Decoders
Control Memory

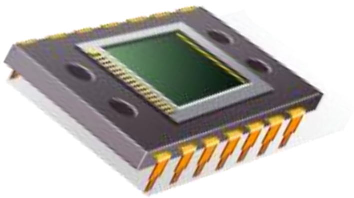
There are four main structural components of the computer:

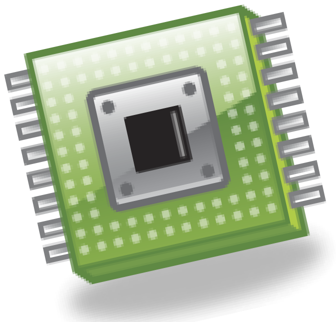


- CPU – controls the operation of the computer and performs its data processing functions
- Main Memory – stores data
- I/O – moves data between the computer and its external environment
- System Interconnection – some mechanism that provides for communication among CPU, main memory, and I/O

# CPU

Major structural components:





- Control Unit
  - Controls the operation of the CPU and hence the computer
- Arithmetic and Logic Unit (ALU)
  - Performs the computer’s data processing function
- Registers
  - Provide storage internal to the CPU
- CPU Interconnection
  - Some mechanism that provides for communication among the control unit, ALU, and registers

# Computer Architecture

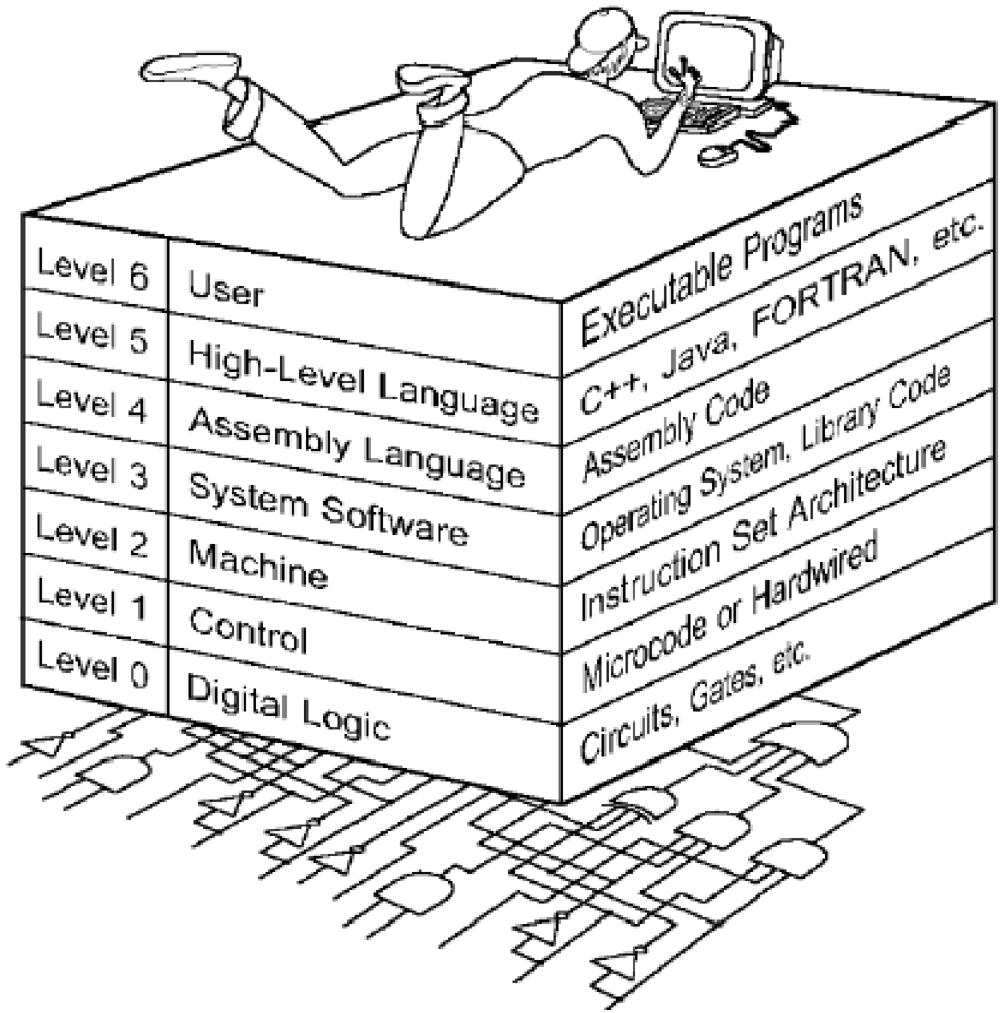
Below your program

# Below your program

- We constantly interact with these computers
  - E.g., via apps on iPhone or via a Word Processor

- How does an application interact with the hardware?

# Computer Level Hierarchy



# Language of the hardware

- The hardware understands *on* or *off*. Hence, the language of the hardware needs two symbols 0 and 1.

- Commonly referred to as binary digits or bits.

- 2 letters do not limit what computers can do (just like the finite letters in our languages)

# Language of the hardware

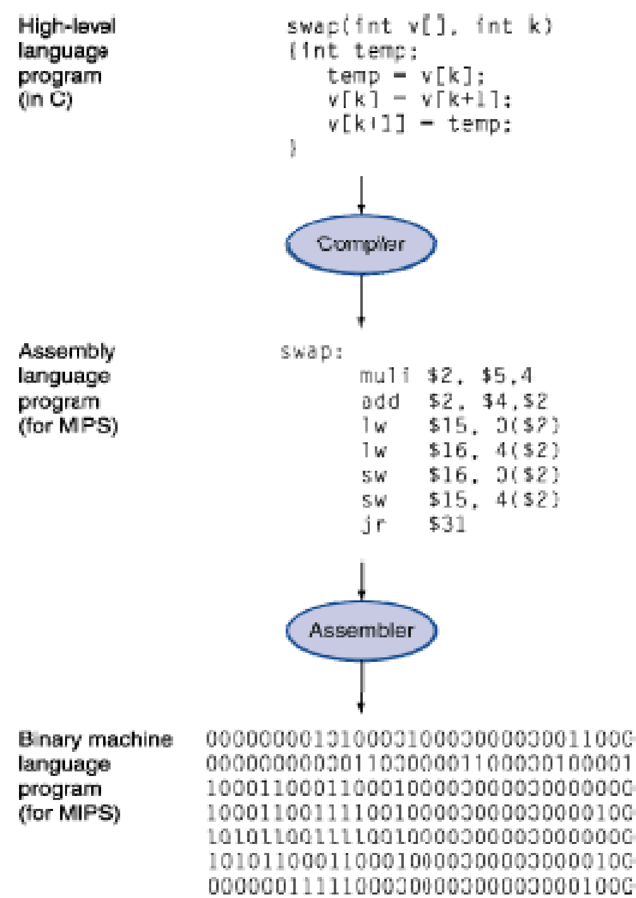
- Instructions are collections of bits that the computer understands

- What our programs instruct the computer to do are the *instructions*:

- Hundreds/thousands/millions of lines of code (instructions) are hidden beneath our apps and programs

# Hierarchical Layers of Program Code

- High-level language
  - Level of abstraction closer to problem domain
  - Provides for productivity and portability
- Assembly language
  - Textual representation of instructions
- Hardware representation
  - Binary digits (bits)
  - Encoded instructions and data



# More on Evolution

# + Semiconductor Memory

**In 1970 Fairchild produced the first relatively capacious semiconductor memory**

- Chip was about the size of a single core
- Could hold 256 bits of memory
- Non-destructive
- Much faster than core

**In 1974 the price per bit of semiconductor memory dropped below the price per bit of core memory**

- There has been a continuing and rapid decline in memory cost accompanied by a corresponding increase in physical memory density
- Developments in memory and processor technologies changed the nature of computers in less than a decade

**Since 1970 semiconductor memory has been through 13 generations**

- Each generation has provided four times the storage density of the previous generation, accompanied by declining cost per bit and declining access time

# Microprocessors

- The density of elements on processor chips continued to rise
  - More and more elements were placed on each chip so that fewer and fewer chips were needed to construct a single computer processor
- 1971 Intel developed 4004
  - First chip to contain all of the components of a CPU on a single chip
  - Birth of microprocessor
- 1972 Intel developed 8008
  - First 8-bit microprocessor
- 1974 Intel developed 8080
  - First general purpose microprocessor
  - Faster, has a richer instruction set, has a large addressing capability

# Evolution of Intel Microprocessors

| | 4004 | 8008 | 8080 | 8086 | 8088 |
|---|---|---|---|---|---|
| Introduced | 1971 | 1972 | 1974 | 1978 | 1979 |
| Clock speeds | 108 kHz | 108 kHz | 2 MHz | 5 MHz, 8 MHz, 10 MHz | 5 MHz, 8 MHz |
| Bus width | 4 bits | 8 bits | 8 bits | 16 bits | 8 bits |
| Number of transistors | 2,300 | 3,500 | 6,000 | 29,000 | 29,000 |
| Feature size ($\mu$m) | 10 | | 6 | 3 | 6 |
| Addressable memory | 640 Bytes | 16 KB | 64 KB | 1 MB | 1 MB |

a. 1970s Processors

| | 80286 | 386TM DX | 386TM SX | 486TM DX CPU |
|---|---|---|---|---|
| Introduced | 1982 | 1985 | 1988 | 1989 |
| Clock speeds | 6 MHz - 12.5 MHz | 16 MHz - 33 MHz | 16 MHz - 33 MHz | 25 MHz - 50 MHz |
| Bus width | 16 bits | 32 bits | 16 bits | 32 bits |
| Number of transistors | 134,000 | 275,000 | 275,000 | 1.2 million |
| Feature size (µm) | 1.5 | 1 | 1 | 0.8 - 1 |
| Addressable memory | 16 MB | 4 GB | 16 MB | 4 GB |
| Virtual memory | 1 GB | 64 TB | 64 TB | 64 TB |
| Cache | — | — | — | 8 kB |

b. 1980s Processors

# Evolution of Intel Microprocessors

| | 486TM SX | Pentium | Pentium Pro | Pentium II |
|---|---|---|---|---|
| Introduced | 1991 | 1993 | 1995 | 1997 |
| Clock speeds | 16 MHz - 33 MHz | 60 MHz - 166 MHz, | 150 MHz - 200 MHz | 200 MHz - 300 MHz |
| Bus width | 32 bits | 32 bits | 64 bits | 64 bits |
| Number of transistors | 1.185 million | 3.1 million | 5.5 million | 7.5 million |
| Feature size ($\mu$m) | 1 | 0.8 | 0.6 | 0.35 |
| Addressable memory | 4 GB | 4 GB | 64 GB | 64 GB |
| Virtual memory | 64 TB | 64 TB | 64 TB | 64 TB |
| Cache | 8 kB | 8 kB | 512 kB L1 and 1 MB L2 | 512 kB L2 |

c. 1990s Processors

| | Pentium III | Pentium 4 | Core 2 Duo | Core i7 EE 990 |
|---|---|---|---|---|
| Introduced | 1999 | 2000 | 2006 | 2011 |
| Clock speeds | 450 - 660 MHz | 1.3 - 1.8 GHz | 1.06 - 1.2 GHz | 3.5 GHz |
| Bus width | 64 bits | 64 bits | 64 bits | 64 bits |
| Number of transistors | 9.5 million | 42 million | 167 million | 1170 million |
| Feature size (nm) | 250 | 180 | 65 | 32 |
| Addressable memory | 64 GB | 64 GB | 64 GB | 64 GB |
| Virtual memory | 64 TB | 64 TB | 64 TB | 64 TB |
| Cache | 512 kB L2 | 256 kB L2 | 2 MB L2 | 1.5 MB L2/12 MB L3 |

d. Recent Processors

# + Microprocessor Speed

Techniques built into contemporary processors include:

## Pipelining

- Processor moves data or instructions into a conceptual pipe with all stages of the pipe processing simultaneously

## Branch prediction

- Processor looks ahead in the instruction code fetched from memory and predicts which branches, or groups of instructions, are likely to be processed next

## Data flow analysis

- Processor analyzes which instructions are dependent on each other's results, or data, to create an optimized schedule of instructions

## Speculative execution

- Using branch prediction and data flow analysis, some processors speculatively execute instructions ahead of their actual appearance in the program execution, holding the results in temporary locations, keeping execution engines as busy as possible

+

# Performance Balance

- Adjust the organization and architecture to compensate for the mismatch among the capabilities of the various components
- Architectural examples include:

**Increase the number of bits that are retrieved at one time by making DRAMs "wider" rather than "deeper" and by using wide bus data paths**

**Reduce the frequency of memory access by incorporating increasingly complex and efficient cache structures between the processor and main memory**

**Change the DRAM interface to make it more efficient by including a cache or other buffering scheme on the DRAM chip**

**Increase the interconnect bandwidth between processors and memory by using higher speed buses and a hierarchy of buses to buffer and structure data flow**

+

# Improvements in Chip Organization and Architecture

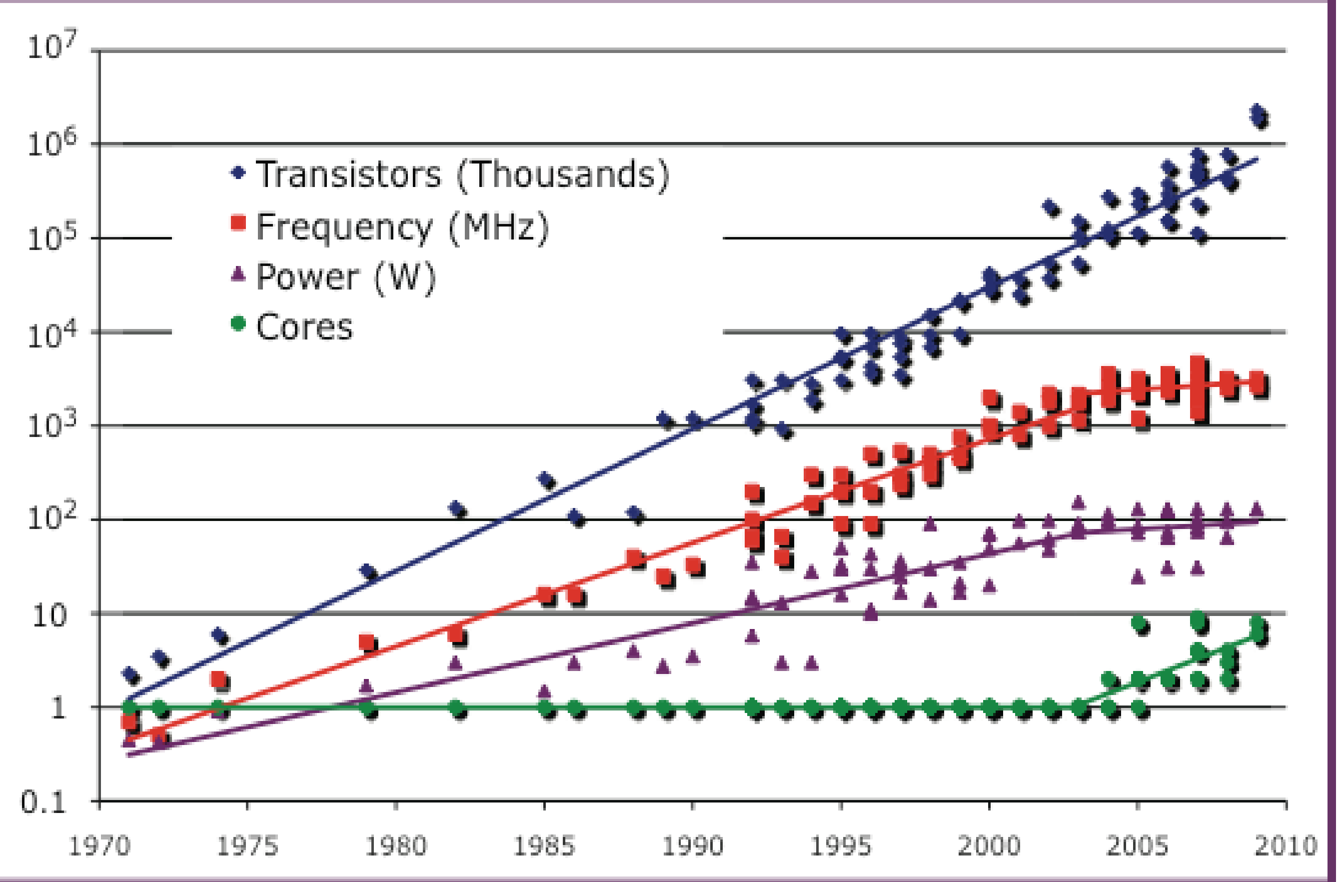
- **Increase hardware speed of processor**
  - Fundamentally due to shrinking logic gate size
    - More gates, packed more tightly, increasing clock rate
    - Propagation time for signals reduced
- **Increase size and speed of caches**
  - Dedicating part of processor chip
    - Cache access times drop significantly
- **Change processor organization and architecture**
  - Increase effective speed of instruction execution
  - Parallelism

# Problems with Clock Speed and Login Density

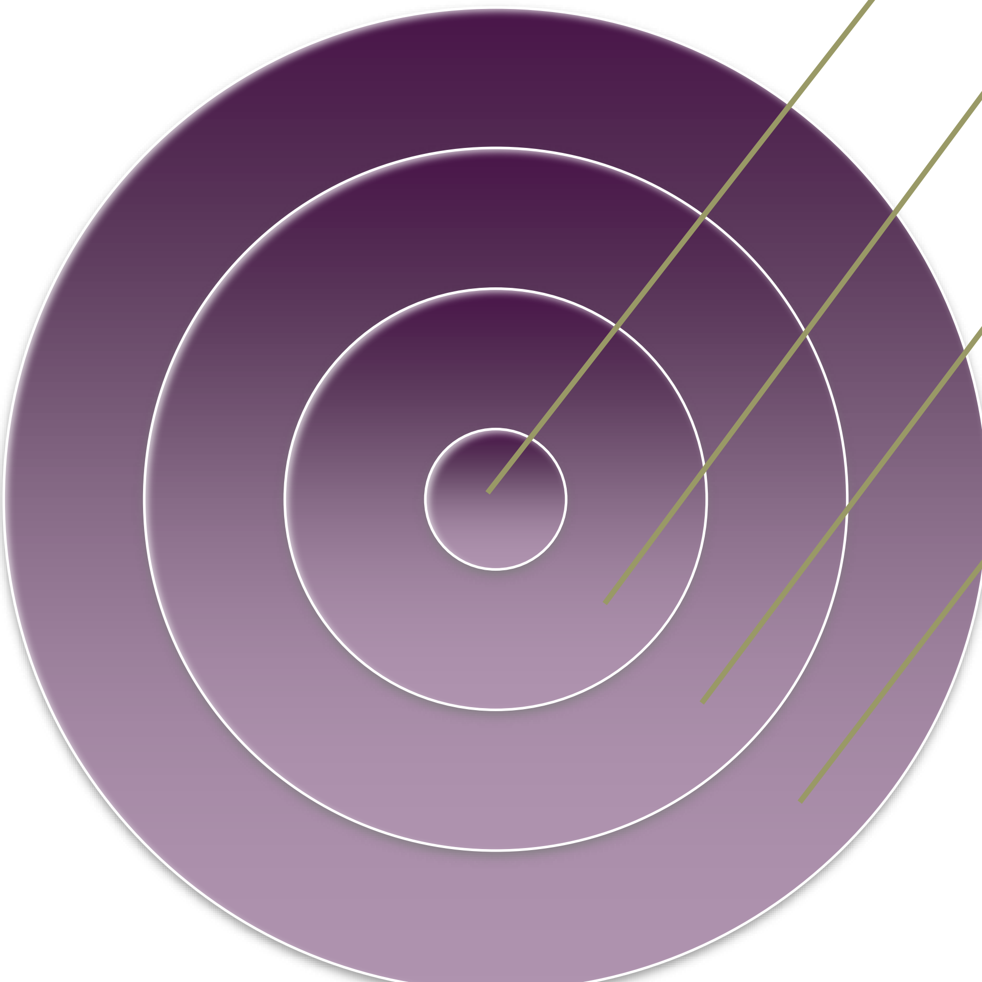
- Power
  - Power density increases with density of logic and clock speed
  - Dissipating heat
- RC delay
  - Speed at which electrons flow on a chip between transistors is limited by resistance and capacitance of metal wires connecting them
  - Delay increases as RC product increases
  - Wire interconnects thinner, increasing resistance
  - Wires closer together, increasing capacitance
- Memory latency
  - Memory speeds lag processor speeds

+

# Processor Trends



# Multicore



- The use of multiple processors on the same chip provides the potential to increase performance without increasing the clock rate
- Strategy is to use two simpler processors on the chip rather than one more complex processor
- With two processors larger caches are justified
- As caches became larger it made performance sense to create two and then three levels of cache on a chip

+

# Many Integrated Core (MIC) Graphics Processing Unit (GPU)

| MIC | GPU |
|---|---|
| ▪ Leap in performance as well as the challenges in developing software to exploit such a large number of cores<br><br>▪ The multicore and MIC strategy involves a homogeneous collection of general purpose processors on a single chip | ▪ Core designed to perform parallel operations on graphics data<br><br>▪ Traditionally found on a plug-in graphics card, it is used to encode and render 2D and 3D graphics as well as process video<br><br>▪ Used as vector processors for a variety of applications that require repetitive computations |

# + Overview

- Results of decades of design effort on complex instruction set computers (CISCs)
- Excellent example of CISC design
- Incorporates the sophisticated design principles once found only on mainframes and supercomputers
- An alternative approach to processor design is the reduced instruction set computer (RISC)
- The ARM architecture is used in a wide variety of embedded systems and is one of the most powerful and best designed RISC based systems on the market
- In terms of market share Intel is ranked as the number one maker of microprocessors for non-embedded systems

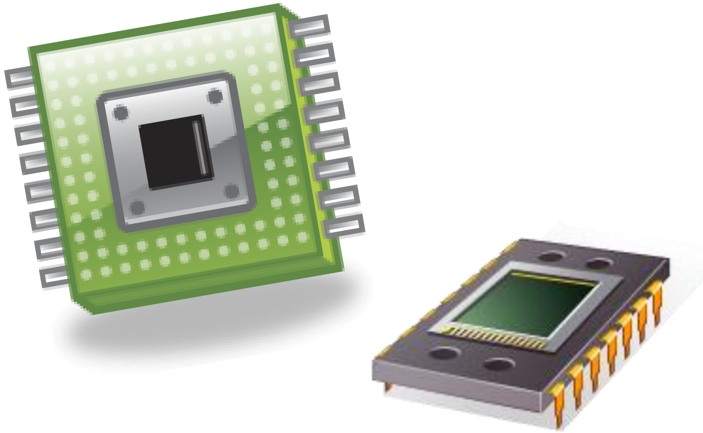
ARM

Intel

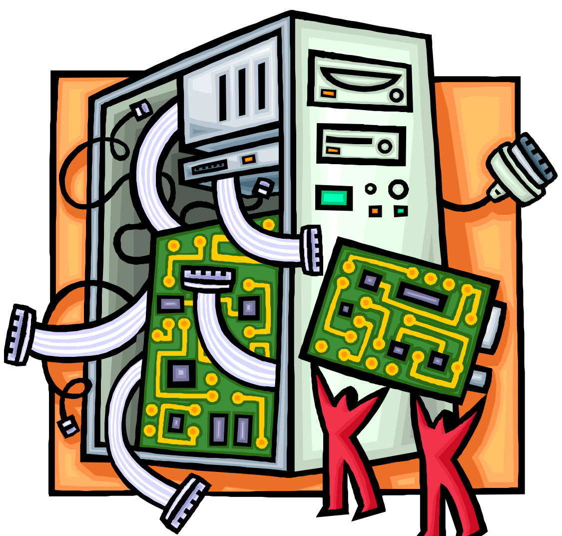
## x86 Architecture

CISC

RISC



# x86 Evolution



8080

- First general purpose microprocessor
- 8-bit machine with an 8-bit data path to memory
- Used in the first personal computer (Altair)

8086

- 16-bit machine
- Used an instruction cache, or queue
- First appearance of the x86 architecture

8088

- used in IBM's first personal computer

80286

- Enabled addressing a 16-MByte memory instead of just 1 MByte

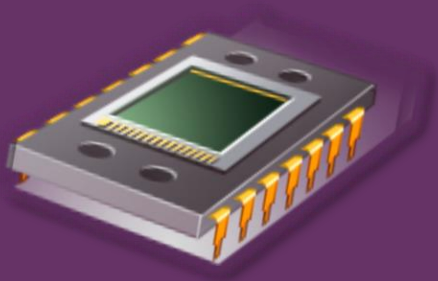
80386

- Intel's first 32-bit machine
- First Intel processor to support multitasking

80486

- More sophisticated cache technology and instruction pipelining
- Built-in math coprocessor

# x86 Evolution - Pentium



**Pentium**

- Superscalar
- Multiple instructions executed in parallel

**Pentium Pro**

- Increased superscalar organization
- Aggressive register renaming
- Branch prediction
- Data flow analysis
- Speculative execution

**Pentium II**

- MMX technology
- Designed specifically to process video, audio, and graphics data

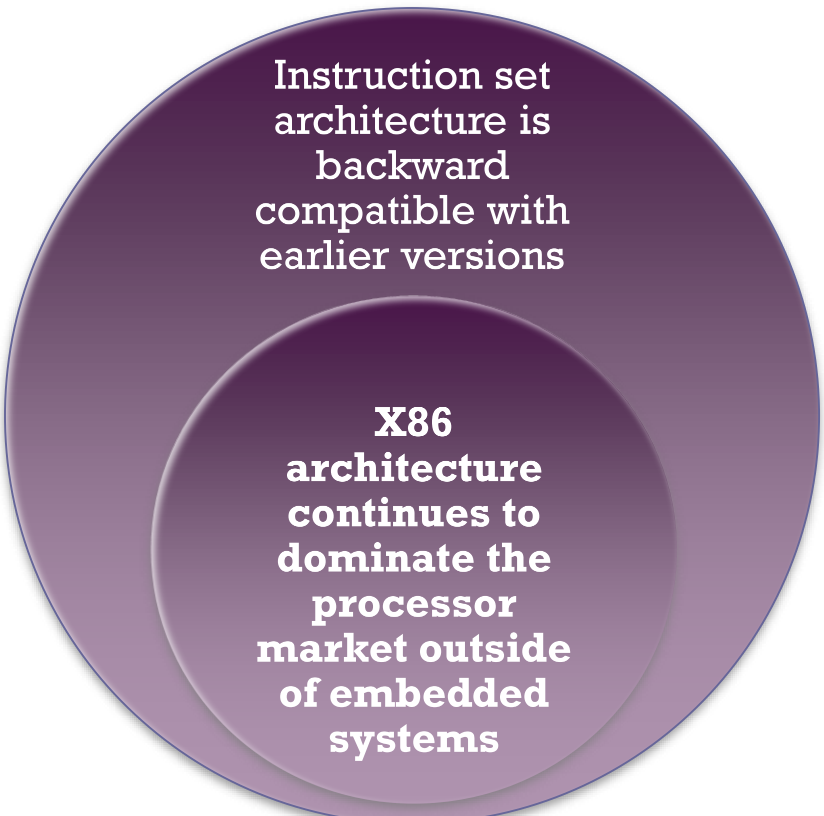
**Pentium III**

- Additional floating-point instructions to support 3D graphics software

**Pentium 4**

- Includes additional floating-point and other enhancements for multimedia

# x86 Evolution (continued)




- Core
  - First Intel x86 microprocessor with a dual core, referring to the implementation of two processors on a single chip
- Core 2
  - Extends the architecture to 64 bits
  - Recent Core offerings have up to 10 processors per chip

Embedded

Systems

Embedded system:

"A combination of computer hardware and software, and perhaps additional mechanical or other parts, designed to perform a dedicated function. In many cases, embedded systems are part of a larger system or product, as in the case of an antilock braking system in a car."

# Examples of Embedded Systems and Their Markets

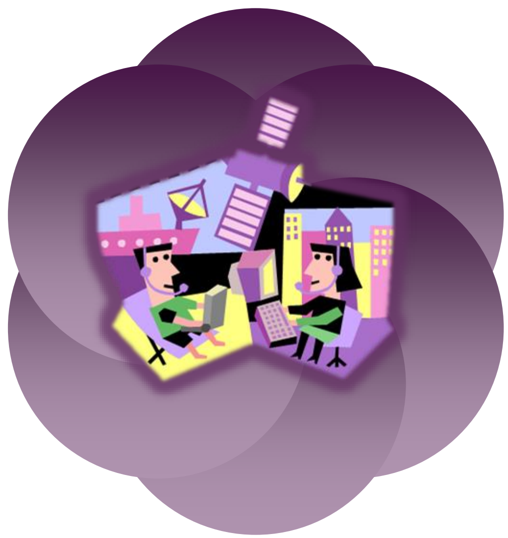
| Market | Embedded Device |
|---|---|
| Automotive | Ignition system<br>Engine control<br>Brake system |
| Consumer electronics | Digital and analog televisions<br>Set-top boxes (DVDs, VCRs, Cable boxes)<br>Personal digital assistants (PDAs)<br>Kitchen appliances (refrigerators, toasters, microwave ovens)<br>Automobiles<br>Toys/games<br>Telephones/cell phones/pagers<br>Cameras<br>Global positioning systems |
| Industrial control | Robotics and controls systems for manufacturing<br>Sensors |
| Medical | Infusion pumps<br>Dialysis machines<br>Prosthetic devices<br>Cardiac monitors |
| Office automation | Fax machine<br>Photocopier<br>Printers<br>Monitors<br>Scanners |

+
# Embedded Systems
## Requirements and Constraints

Small to large systems, implying different cost constraints and different needs for optimization and reuse

Relaxed to very strict requirements and combinations of different quality requirements with respect to safety, reliability, real-time and flexibility

Different models of computation ranging from discrete event systems to hybrid systems



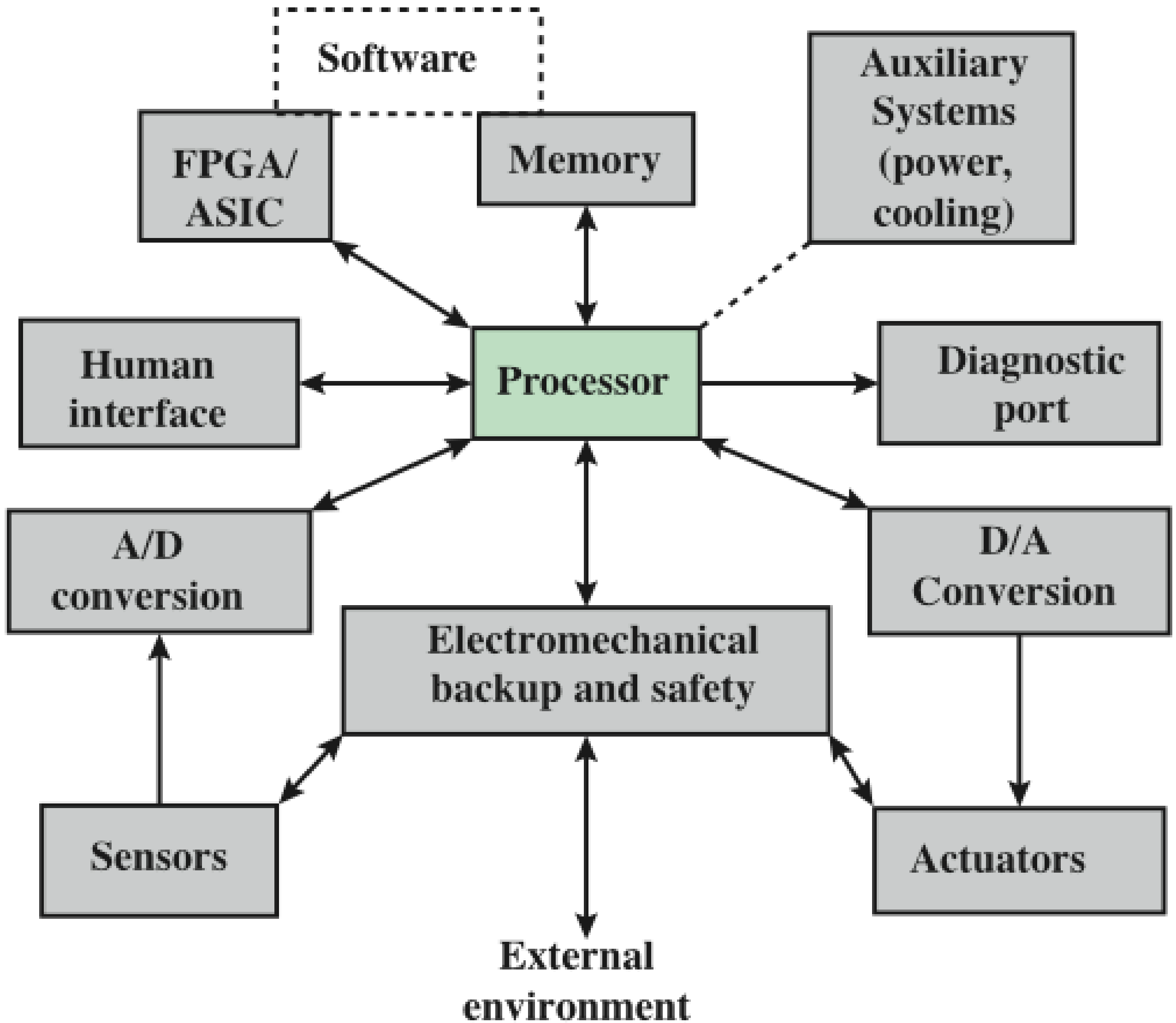
Short to long life times

Different application characteristics resulting in static versus dynamic loads, slow to fast speed, compute versus interface intensive tasks, and/or combinations thereof

Different environmental conditions in terms of radiation, vibrations, and humidity

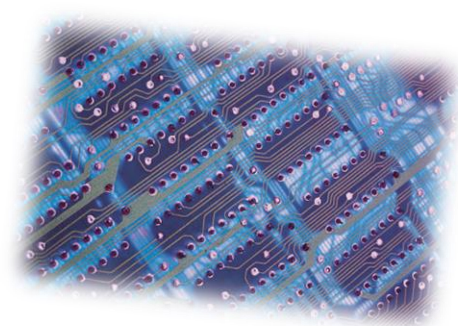
+

# Possible Organization of an Embedded System



# Acorn RISC Machine (ARM)

- Family of RISC-based microprocessors and microcontrollers
- Designs microprocessor and multicore architectures and licenses them to manufacturers
- Chips are high-speed processors that are known for their small die size and low power requirements
- Widely used in PDAs and other handheld devices
- Chips are the processors in iPod and iPhone devices
- Most widely used embedded processor architecture
- Most widely used processor architecture of any kind



# ARM Evolution

| Family | Notable Features | Cache | Typical MIPS @ MHz |
|---|---|---|---|
| ARM1 | 32-bit RISC | None | |
| ARM2 | Multiply and swap instructions; Integrated memory management unit, graphics and I/O processor | None | 7 MIPS @ 12 MHz |
| ARM3 | First use of processor cache | 4 KB unified | 12 MIPS @ 25 MHz |
| ARM6 | First to support 32-bit addresses; floating-point unit | 4 KB unified | 28 MIPS @ 33 MHz |
| ARM7 | Integrated SoC | 8 KB unified | 60 MIPS @ 60 MHz |
| ARM8 | 5-stage pipeline; static branch prediction | 8 KB unified | 84 MIPS @ 72 MHz |
| ARM9 | | 16 KB/16 KB | 300 MIPS @ 300 MHz |
| ARM9E | Enhanced DSP instructions | 16 KB/16 KB | 220 MIPS @ 200 MHz |
| ARM10E | 6-stage pipeline | 32 KB/32 KB | |
| ARM11 | 9-stage pipeline | Variable | 740 MIPS @ 665 MHz |
| Cortex | 13-stage superscalar pipeline | Variable | 2000 MIPS @ 1 GHz |
| XScale | Applications processor; 7-stage pipeline | 32 KB/32 KB L1<br>512 KB L2 | 1000 MIPS @ 1.25 GHz |

DSP = digital signal processor

SoC = system on a chip

# ARM Design Categories

- ARM processors are designed to meet the needs of three system categories:

- Secure applications
  - Smart cards, SIM cards, and payment terminals

- Embedded real-time systems
  - Systems for storage, automotive body and power-train, industrial, and networking applications

- Application platforms
  - Devices running open operating systems including Linux, Palm OS, Symbian OS, and Windows CE in wireless, consumer entertainment and digital imaging applications