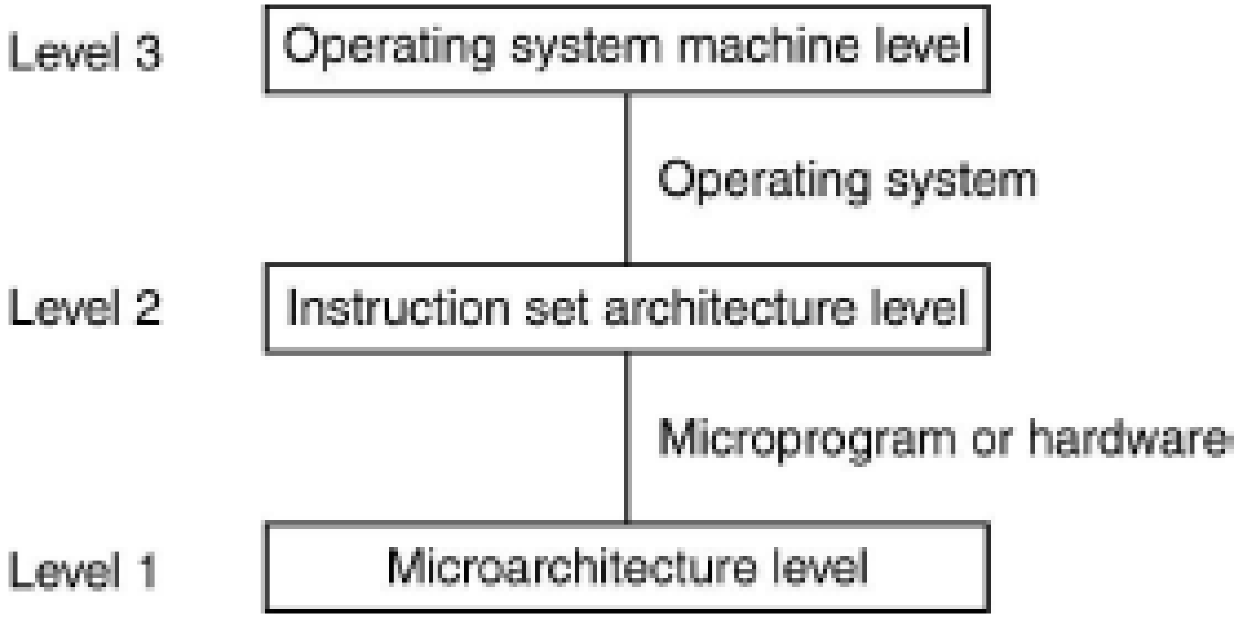
CAO: Lecture 11
Multilevel Viewpoint II

Topics Covered

- Operating System Machine
- What is an OS
- Goal of OS
- Instruction Set Representation
- Instruction Formats
- Registers
- CPU
- ALU
- Control Unit
- CPU

Operating System Machine



Positioning of the operating system machine level.

Introduction

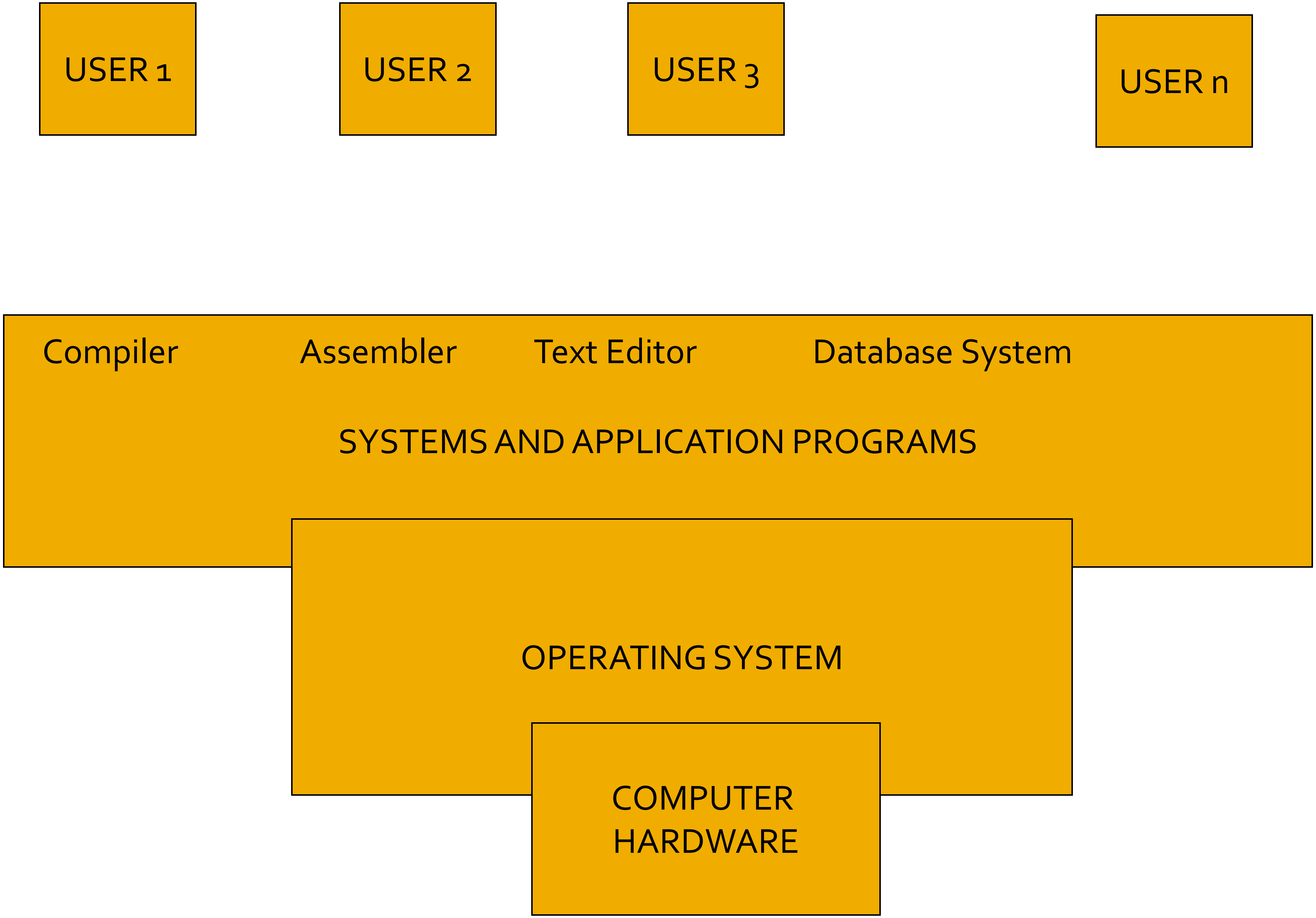
- An operating system is the part of the program that manages the computer hardware.
- It provides basis for application programs and acts as an intermediary between user of a computer and computer hardware.

- OS varies in accomplishing tasks.........
- For example:
  - Mainframe OS are designed primarily to optimize utilization of hardware.
  - P4ersonal computer operating system support complex and business application and everything in between.
  - Handheld computer OS are deigned to provide an environment in which a user can easily interface with the computer to execute programs.

Thus some OS are designed to be convenient, others to be efficient, and other some combination of two.

What is an OS

- A computer can be divided into four componenets:
  - Hardware
  - Operating system
  - Application programs
  - Users



ABSTRACT VIEW OF THE COMPONENETS OF THE COMPUTER SYSTEM

- **The hardware –** the **central processing unit,** the **memory, input/output -** provides the basic computing resources**.**

- The application resources – such as word processors, spreadsheet, compilers and web browsers – define the ways in which these resources are used to solve computing problems of the users.

- The OS controls and coordinates the use of hardware among various application programs for the various users

OS can be viewed from two viewpoints :

- Users view
- Systems view

Users view

- **User1**: sits in front of PC, consisting of a monitor, keyboard, mouse and system unit.
  - Such a system is designed for one user to monopolize its resources, to maximize the work (or play) that the user is performing
  - In this case, the OS is designed mostly for **ease of use,** with some attention paid to performance and none paid to resource utilization.
  - Performance is important to the user, but it does not matter if most system is sitting idle, waiting for the slow I/O speed of the user

- **User2:** Some users sit at a terminal connected to a mainframe or minicomputer. Other users are accessing the same computer through other terminals.
  - These users share resources and exchange information
  - OS is designed to maximize resource utilization – to assure that all available CPU time, memory and I/O are used efficiently.

Users view contd……..

- **<u>User3</u>**: sit at workstation connected to networks of other workstations and servers. These users have dedicated resources at their disposal, but they also share resources such as networking and servers – file, compute and print servers. Therefore OS is designed to compromise between individual usability and resource utilization.
- **<u>User4</u>**: handheld computers – individual users – wireless . OS is designed for individual usability but performance per battery life is important.

- **<u>User 5</u>**:Some computers have little or no user intervention. Embedded computers at home and automobile may have numeric keypad and may turn indicator lights On or OFF to show status but mostly they and their OS are deigned to run without user intervention

Systems View

- Os is the most intimate with the hardware
- We can view it as a **RESOURCE ALLOCATOR** -----

- COMPUTER System has many resources – h/w and s/w – they may be required to solve a problem : CPU time, memory space, file-storage space, I/O devices and so on….

- OS acts as an manger of these resources, the OS must decide how to allocate them to specific programs and users so that it can operate the computer system efficiently.

Systems View Contd.....

- OS emphasize the need to control various I/O devices and user programs.
- Os is a control program which manages the execution of the user programs to prevent errors and improper use of the computer. ----especially concerned with operation and control of I/O devices.

OS is one program which is running at all times on the computer

Goal of OS

- Convenient for the user
- Efficient in use.

OS supports

- Job scheduling
- Multiprogramming – os sits never idle
- Multitasking
- Deadlock prevention

- An **operation system** is a program that adds a variety of new instructions and features, above and beyond what the ISA level provides.
- OSM (Operating System Machine) level instructions are available to applications programmers. It contains all ISA level instructions and the set of new instructions.

.

- Multiprogramming : one processor …multiple programs…….one user….. …ex: windows

- Multitasking: one processor ……. Multiple programs …….. Multiple user…ex: Unix

Instruction Set Architecture (ISA)

The Instruction Set Architecture (ISA) is the part of the processor that is visible to the programmer or compiler writer. The ISA serves as the boundary between software and hardware.

The ISA can be described using 5 categories:

- Operand Storage in the CPU
- Where are the operands kept other than in memory?
- Number of explicit named operands
- How many operands are named in a typical instruction.
- Operand location
- Can any ALU instruction operand be located in memory? Or must all operands be kept internally in the CPU?
- Operations
- What operations are provided in the ISA.
- Type and size of operands
- What is the type and size of each operand and how is it specified?

The 3 most common types of ISAs are:

Stack - The operands are implicitly on top of the stack.

Accumulator - One operand is implicitly the accumulator.

General Purpose Register (GPR) - All operands are explicitely mentioned, they are either registers or memory locations.

```
A = B + C;
```

in all 3 architectures:

Stack	Accumulator	GPR
PUSH A	LOAD A	LOAD R1,A
PUSH B	ADD B	ADD R1,B
ADD	STORE C	STORE R1,C
POP C	-	-

Instruction Set Representation

- Purpose of instruction representation
  - Conversion of high level language program and data structures into machine level instructions
- Instruction set architecture
  - Interface between high level language and machine language
  - Instruction set
  - Addressing modes
  - Instruction format

High Level Language

- Data types and structures:
  - Scalars – Integers, Real numbers, Characters
  - Arrays
  - Records (structs)
  - Arrays of records
  - Pointers
- Program structures:
  - Assignment statements
  - Arithmetic expressions
  - Control transfer statements
    - If-then-else statement
    - Switch statement
    - Loop statements
    - Function call statement
    - Function return statement
  - File I/O statements

Machine Language

- Instructions
  - Logical instructions
    - AND, OR, XOR, Shift
  - Arithmetic instructions
    - Data types
      - Integers: Unsigned, Signed, Byte, Short, Long
      - Real numbers: Single-precision (float), Double-precision (double)
    - Operations
      - Addition, Subtraction, Multiplication, Division
  - Data transfer instructions
    - Register transfer: Move
    - Memory transfer: Load, Store
    - I/O transfer: In, Out
  - Control transfer instructions
    - Unconditional branch
    - Conditional branch
    - Procedure call
    - Return

Instruction Formats

- **3-operand instructions**
  - ADD op1, op2, op3; op1 ⟵ op2 + op3
- **2-operand instructions**
  - ADD op1, op2; op1 ⟵ op1 + op2
- **1-operand instructions**
  - INC op1; op1 ⟵ op1 + 1
- **Types of operands:**
  - Register operands
  - Memory operands specified using addressing modes

Instruction Set Architectures

- **Complex Instruction Set Computer (CISC) processors:**
  - 2-operand instructions and 1-operand instructions
  - Any instruction can use memory operands
  - Many addressing modes
  - Complex instruction formats: Varying length instructions

- **Reduced Instruction Set Computer (RISC) processors:**
  - 3-operand instructions, 2-operand instructions, and 1-operand instructions
  - Load-Store Architecture (LSA) processors:
    - Only memory transfer instructions (Load and Store) can use memory operands.
    - All other instructions can use register operands only.
  - Simple instruction formats: Fixed length instructions

Translation of High-Level Language Statements

Assignment statement: A = B + C

- CISC Architecture with 3-operand instructions:

```
ADD     A, B, C
```

Opcode of ADD	Address of A	Address of B	Address of C

- CISC Architecture with 2-operand instructions:

```
LOAD    R0, B
ADD     R0, C
STORE   A, R0
```

Opcode of ADD	Register R0	Address of C

- RISC Architecture (Load-Store Architecture):

```
LOAD    R0, B
LOAD    R1, C
ADD     R2, R0, R1
STORE   A, R2
```

Opcode of ADD	Register R2	Register R0	Register R1

Registers

- The **registers set** stores intermediate data used during the execution of the instructions

- **Processor register** is a small amount of very fast computer memory used to speed the execution of computer programs by providing quick access to commonly used values—typically, the values being in the midst of a calculation at a given point in time. (AC)

- **Data registers** are used to store integer numbers (DR)

- **Address registers** hold memory addresses and are used to access memory. (AR)

- **General Purpose registers** (**GPR**s) can store both data and addresses, i.e., they are combined Data/Address registers

- **Floating Point registers** (**FPR**s) are used to store floating pointComputer arithmetic

- **Constant registers** hold read-only values (e.g., zero, one, pi, ...).

- **Vector registers** hold data for vector processing

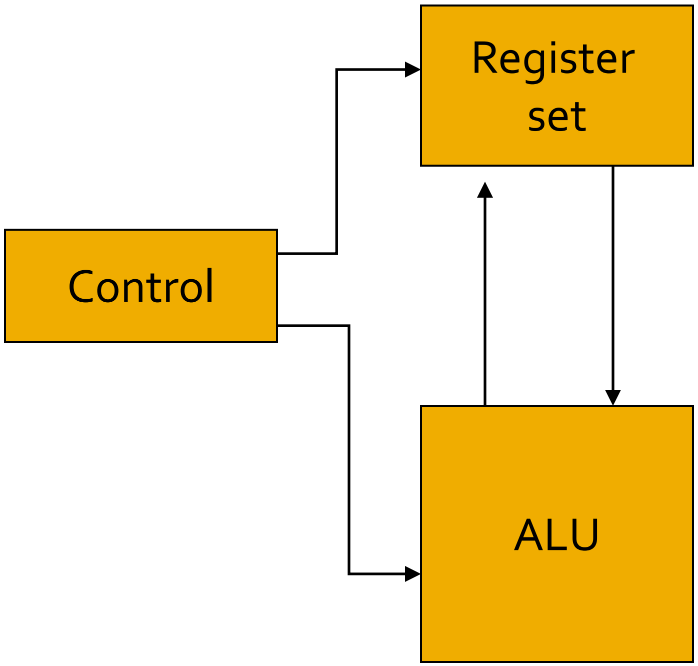
- **Special Purpose registers** store internal CPU data, like the program counter which indicates where the computer is in its instruction sequence

- **Control Registers** which ctrl the general behavior of the CPU

- Program counter: Holds address for instruction (i.e. address of the next instruction after execution of the current instruction is completed (PC)
- Instruction Register: holds the instruction code. (IR)
- Temporary register: holds temporary data (TR)
- Input register: holds input character (INPR)
- Output registers: holds output character (OUTR)

Central Processing Unit

- The part of the computer that performs the bulk of data processing operations is called the central processing unit and is referred to as CPU.
- The **registers set** stores intermediate data used during the execution of the instructions
- The **ALU** performs the required micro operations for executing the instructions.
- The **Control unit** supervises the transfer of information among the registers and instructs the ALU as to which operation to perform.

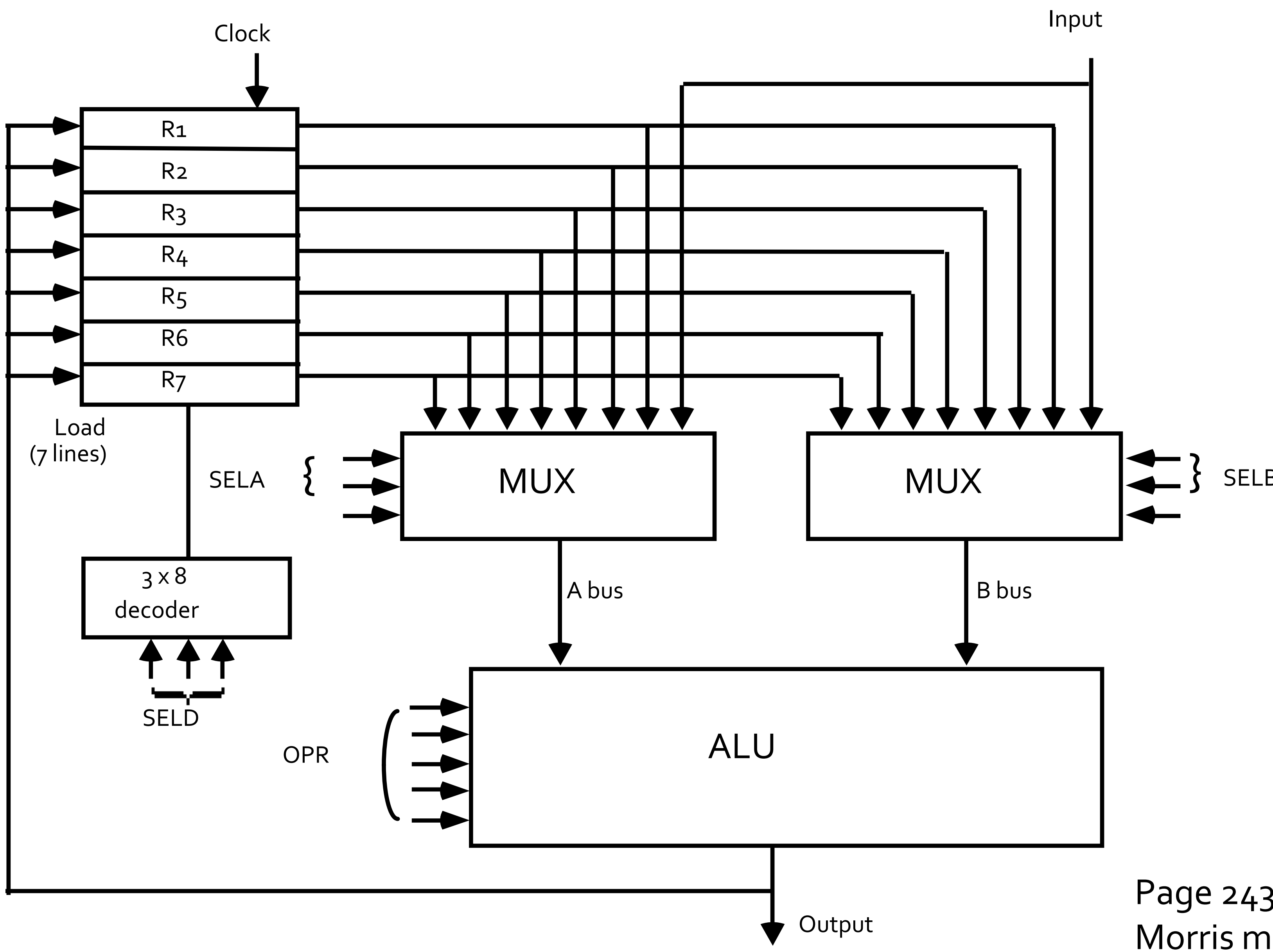


Major Components of CPU

GENERAL REGISTER ORGANIZATION

- Memory locations are needed for storing pointers, Return addresses etc.
- Referring to memory locations for such applications is time consuming because memory access is most time consuming operation in computer.
- So, it is more convenient and efficient to store these intermediate values in processor registers
- When large number of registers are included in the CPU it is efficient to connect them through common bus system.
- Because registers communicate with each other not only for direct data transfers, but also while performing various microoperations

- **What is BUS**
- Dig tal computers has many registers and path must be provided to transfer information from one register to another.
- No. of wires will be excessive if separate lines are used between each register. Most efficient way is to have Common bus system.
- Bus structure consists of a set of common lines, one for each bit of registers, thru which binary information is transferred one at a time.
- Ctrl signals determine which register is selected by the bus during each particular register transfer



Page 243
Morris mano

- **General Register Organization**:—
- When a large number of registers are included in the CPU, it is most efficient to connect them through a common bus system. The registers communicate with each other not only for direct data transfers, but also while performing various micro-operations. Hence it is necessary to provide a common unit that can perform all the arithmetic, logic and shift micro-operation in the processor.

- **A Bus organization for seven CPU registers:—**
- **Reference Diagram: Page Number 243 by M Morris Mano**

- The output of each register is connected to true multiplexer (mux) to form the two buses A & B. The selection lines in each multiplexer select one register or the input data for the particular bus. The A and B buses forms the input to a common ALU. The operation selected in the ALU determines the arithmetic or logic micro-operation that is to be performed. The result of the micro-operation is available for output and also goes into the inputs of the registers. The register that receives the information from the output bus is selected by a decoder. The decoder activates one of the register load inputs, thus providing a transfer both between the data in the output bus and the inputs of the selected destination register.

- The control unit that operates the CPU bus system directs the information flow through the registers and ALU by selecting the various components in the systems.
- R1 → R2 + R3
- (1) MUX A selection (SEC A): to place the content of R2 into bus A
- (2) MUX B selection (sec B): to place the content of R3 into bus B
- (3) ALU operation selection (OPR): to provide the arithmetic addition (A + B)
- (4) Decoder destination selection (SEC D): to transfer the content of the output bus into R1
- These form the control selection variables are generated in the control unit and must be available at the beginning of a clock cycle.

ALU

- **Arithmetic:**
- Addition, Subtraction, Multiplication, Division
- **Logic:**
- Comparisons

Control Unit

- Reads & Interprets Program Instructions
- Directs the Operation of the Processor
- Controls the flow of programs and data into and out of memory

- CU consists of two decoders, a counter and a number of ctrl logic gates.

- An instruction read from memory is placed in instruction register (IR) where it is divided into three parts:
- I bit
- Opcode
- 0-11 bits ----

- Operation code in bits 12 thru 14 are decoded with a 3*8 decoder

Central Processing Unit

- Machine Cycle
- Fetch
- Decode
- Execute
- Store

- Fetch
- Calls an instruction into memory

- Decode
- Figures out what the instruction is trying to do

- Execute
- Does the decoded instruction
- Add 2+2

- Store
- Puts the answer 4 into memory for use by another instruction

Memory

- Memory unit is needed for storing programs and data.
- Memory units that communicate directly with CPU is called MAIN MEMORY
- Devices that provide backup storage is called **auxiliary memory**
- Most common auxiliary memory is magnetic disks and magnetic tapes. They are used for storing programs, large data files, and other backup information.
- Only programs and data that are currently needed by the processor will reside in main memory
- All other information is stored in auxiliary memory and transferred to main memory when needed

- **<u>Auxiliary memory</u>**
- Slow
- High Capacity

<

- **<u>Main memory</u>**
- Small but relatively faster than auxiliary **memory**

<

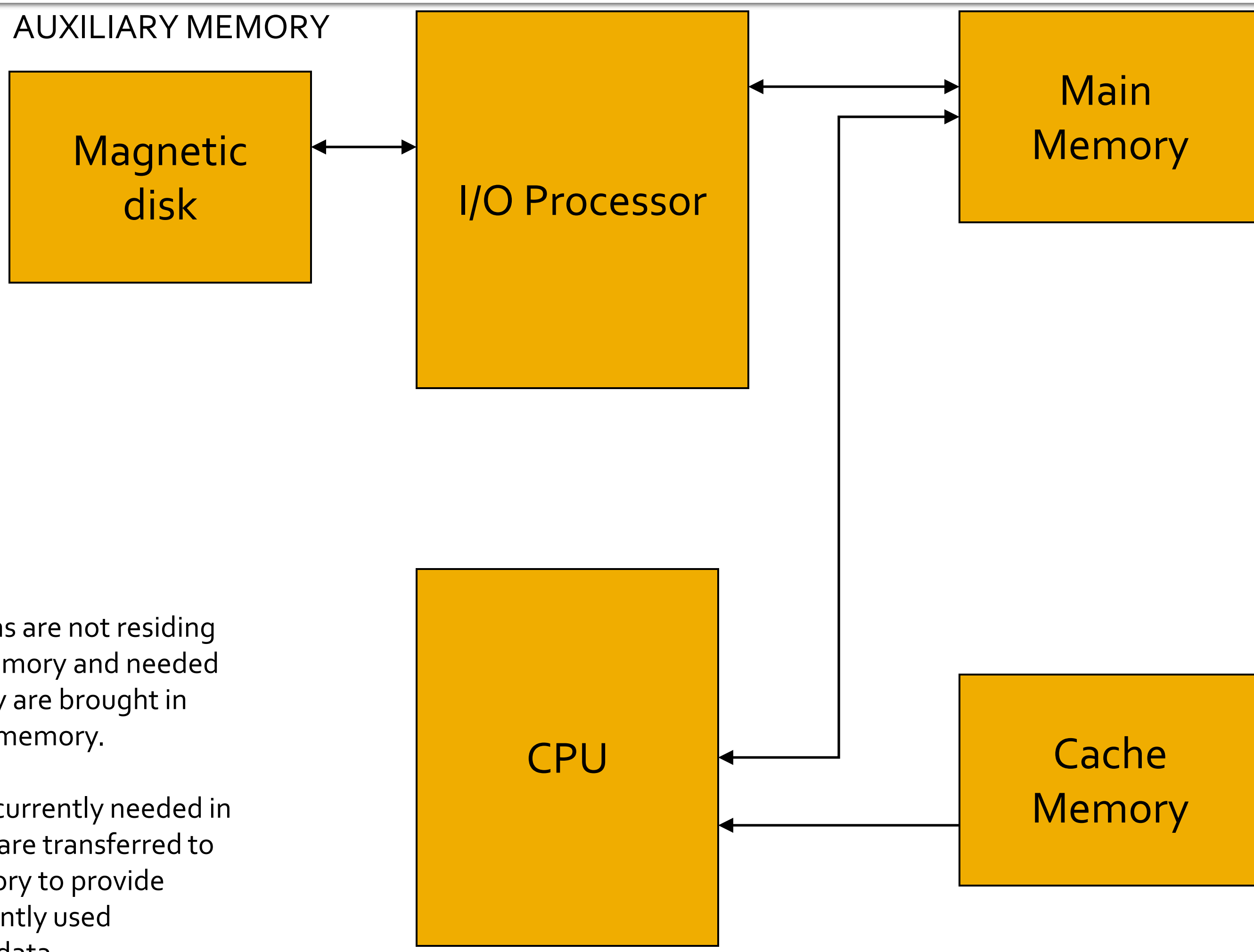
<u>Cache memory</u>

Smaller and faster

Memory hierarchy in computer system

Slow magnetic tapes used to store removable files

Used as backup storage



occupies the central position by being able to communicate directly with the CPU and with auxiliary memory through /]/o processor

When programs are not residing in the main memory and needed by the cpu they are brought in from auxiliary memory.

Programs not currently needed in main memory are transferred to auxiliary memory to provide space for currently used programs and data.

CACHE MEMORY

- a special very high speed memory called is sometimes used to increase the speed of the processing by making current programs and data available to the CPU at the rapid rate.
- The cache memory is employed in computer systems to compensate the speed differential between main memory access time and logic.
- CPU logic is usually faster than main memory access time, with the result that processing speed is limited primarily by the speed of the main memory.
- A Technique used to compensate for the mismatch in operating speeds is to employ an extremely fast, small cache between the CPU and main memory whose access time is close to processor logic clock cycle time.
- The cache is used for storing segments of programs currently being executed in the CPU and temporary data frequently needed in the present calculations
- By doing this the performance rate of the computer also increases

Main Memory / Primary Memory units

- **RAM (Random Access Memory)**
- **ROM (Read-only Memory)**

- – They work in different ways and perform distinct functions
- – CPU Registers
- – Cache Memory

- • Also termed as 'auxiliary' or 'backup' storage, it is typically used as a
- supplement to main storage. It is much cheaper than the main storage and
- stores large amount of data and instructions permanently. Hardware devices
- like magnetic tapes and disks fall under this category.

Secondary Memory/Auxiliary Memory

- Also termed as ‘auxiliary’ or ‘backup’ storage, it is typically used as a supplement to main storage.
- It is much cheaper than the main storage and stores large amount of data and instructions permanently.
- Hardware devices like magnetic tapes and disks fall under this category

Random Access Memory

- RAM or Random Access Memory is the central storage unit in a computer system.
- It is the place in a computer where the operating system, application programs and the data in current use are kept temporarily so that they can be accessed by the computer's processor.
- The more RAM a computer has, the more data a computer can manipulate.
- Random access memory, also called the Read/Write memory, is the temporary memory of a computer.
- It is said to be 'volatile' since its contents are accessible only as long as the computer is on.
- The contents of RAM are cleared once the computer is turned off.

- Types of RAM

1. STATIC RAM
2. DYNAMIC RAM

- **STATIC RAM:** CONSISTS OF INTERNAL FLIP FLOPS THAT STORES THE BINARY INFORMATION . The stored information remains valid as long as the power is applied to the unit.
- **DYNAMIC RAM:** stores the binary information in the form of electric charges that are applied to the capacitors. (capacitors are attached to transistors) The capacitors are provided by the inside the chip by the MOS (metal oxide transistor) transistors. The stored charge on the capacitors tend to discharge with time and the capacitors must be periodically recharged by refreshing the dynamic memory.
- Static memory is easy to use.

ROM

- ROM or Read Only Memory is a special type of memory which can only be read and contents of which are not lost even when the computer is switched off.
- It typically contains manufacturer's instructions.
- Among other things, ROM also stores an initial program called the **'bootstrap loader'** whose function is to start the computer software operating, once the power is turned on.
- Contents of ROM remains unchanged after the power is turned off and on again.

- Read-only memories can be manufacturer-programmed or user-programmed.

- **PROM**
- While **manufacturer-programmed** ROMs have data burnt into the circuitry, **user programmed** ROMs can have the user load and then store read-only programs.

- **EPROM**
- Information once stored on the ROM or PROM chip cannot be altered. However, another type of memory called EPROM (Erasable PROM) allows a user to erase the information stored on the chip and reprogram it with new information.

ROM Types

- PROM
- EPROM
- EEPROM

- Each type has unique characteristics, but they are all types of memory with two things in common:
- Data stored in these chips is nonvolatile -- it is not lost when power is removed.
- Data stored in these chips is either unchangeable or requires a special operation to change (unlike RAM, which can be changed as easily as it is read).

- ROM chips are fundamentally different from RAM chips. While RAM uses transistors to turn on or off access to a capacitor at each intersection, ROM uses a **diode**

AUXILIARY MEMORY

- RAM is volatile memory having a limited storage capacity. Secondary/auxiliary storage is storage other than the RAM.
- These include devices that are peripheral and are connected and controlled by the computer to enable permanent storage of programs and data.
- The memory is specifically meaning the RAM. This keeps the information for a shorter period of time (usually volatile), is faster and more expensive.
- By Storage we mean the Hard disk. Here the information is retained longer (nonvolatile), It's Slower and Cheaper

Auxiliary Storage Devices-Magnetic Tape, Floppy Disk, Hard Disk.

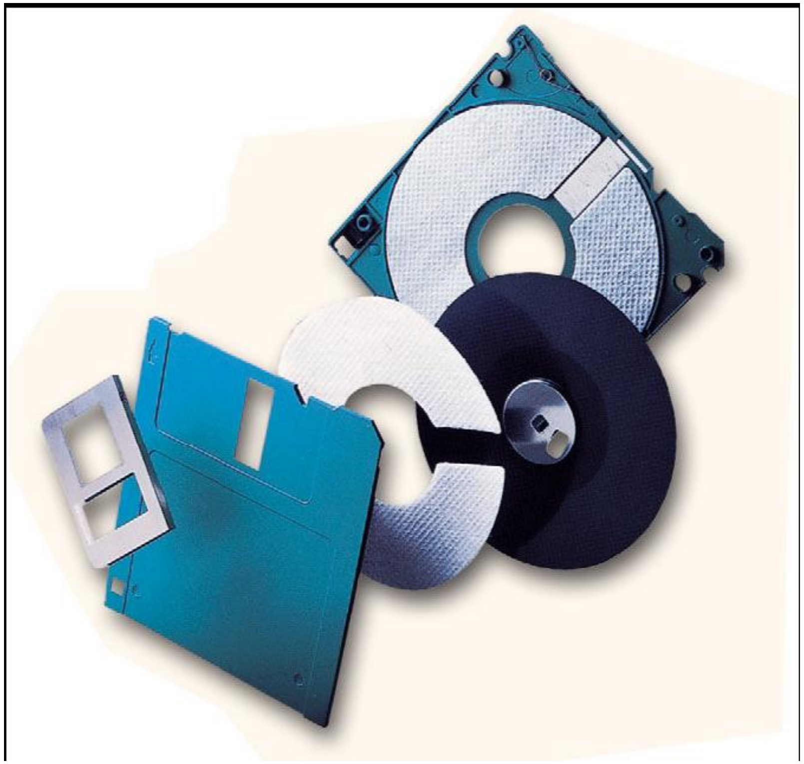
- The Magnetic Storage Exploits duality of magnetism and electricity. It converts electrical signals into magnetic charges, captures magnetic charge on a storagemedium and then later regenerates electrical current from stored magnetic charge.Polarity of magnetic charge represents bit values zero and one.

- **Magnetic Disk**
- The Magnetic Disk is Flat, circular platter with metallic coating that is rotated beneath read/write heads. It is a Random access device; read/write head can be moved to any location on the platter.

Floppy Disk

- These are small removable disks that are plastic coated with magnetic recording material. Floppy disks are typically 3.5" in size (diameter) and can hold 1.44 MB of data. This portable storage device is a rewritable media and can be reused a number of times.

- Floppy disks are commonly used to move files between different computers. The main disadvantage of floppy disks is that they can be damaged easily and, therefore, are not very reliable.



HARD DISK

- Another form of auxiliary storage is a hard disk. A hard disk consists of one or more rigid metal plates coated with a metal oxide material that allows data to be magnetically recorded on the surface of the platters.

- The hard disk platters spin at a high rate of speed, typically 5400 to 7200 revolutions per minute (RPM).

- Storage capacities of hard disks for personal computers range from 10 GB to 120 GB (one billion bytes are called a gigabyte).

Optical Drives

- CD's (Compact Disk)
- Their storage:
- ~ 700 MB storage
- Their Types:
- – CD-ROM (read only)
- – CD-R: (record) to a CD
- – CD-RW: can write and erase CD to reuse it (re-writable)
- – DVD(Digital Video Disk)