JGI JAIN DEEMED-TO-BE UNIVERSITY | SCHOOL OF COMPUTER SCIENCE AND IT

Master of Computer Applications

23MCAC105 – Advanced Computer Architecture

Credits: 3

L: T: P – 3-0-0

Prepared by

Dr. A. Rengarajan, Professor

Module – 3

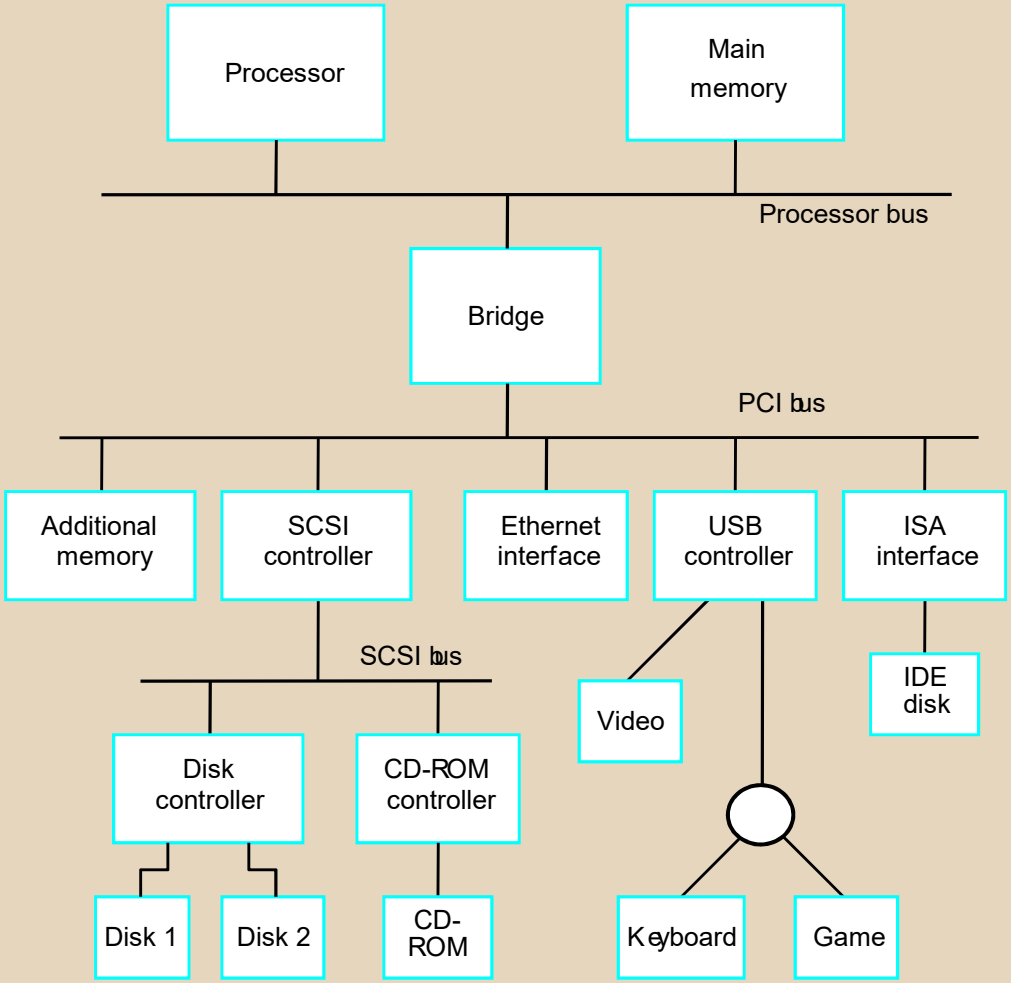
I/O and MEMORY

- I/O Interface
- PCI Bus
- SCSI Bus
- USB
- Data Transfer: Serial, Parallel
- Synchronous, Asynchronous Modes of Data Transfer
- Direct Memory Access (DMA)
- I/O Processor
- Main memory- RAM, ROM
- Secondary Memory – Magnetic Tape, Disk, Optical Storage
- Cache Memory

I/O Interface

- ✓ The needs for standardized interface signals and protocols.
- ✓ Motherboard
- ✓ Bridge: circuit to connect two buses
- ✓ Expansion bus
- ✓ ISA, PCI, SCSI, USB,…

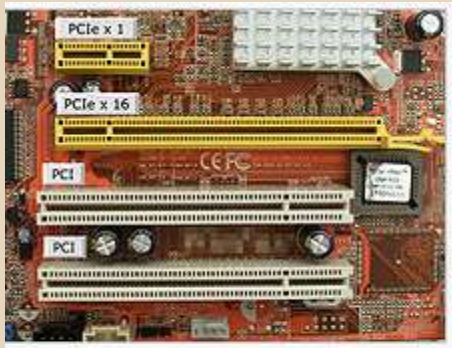
I/O Interface



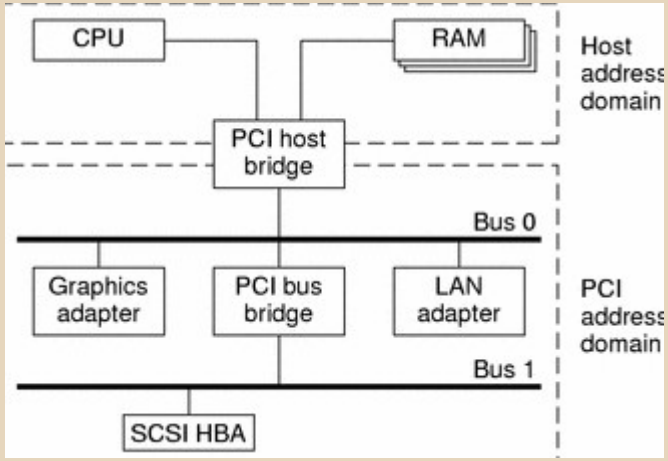
An example of a computer system using different interface standards.

PCI Bus

- ✓ **Peripheral Component Interconnect (PCI)** is a local computer bus for attaching hardware devices in a computer and is part of the PCI Local Bus standard.
- ✓ The PCI bus supports the functions found on a processor bus but in a standardized format that is independent of any given processor's native bus.



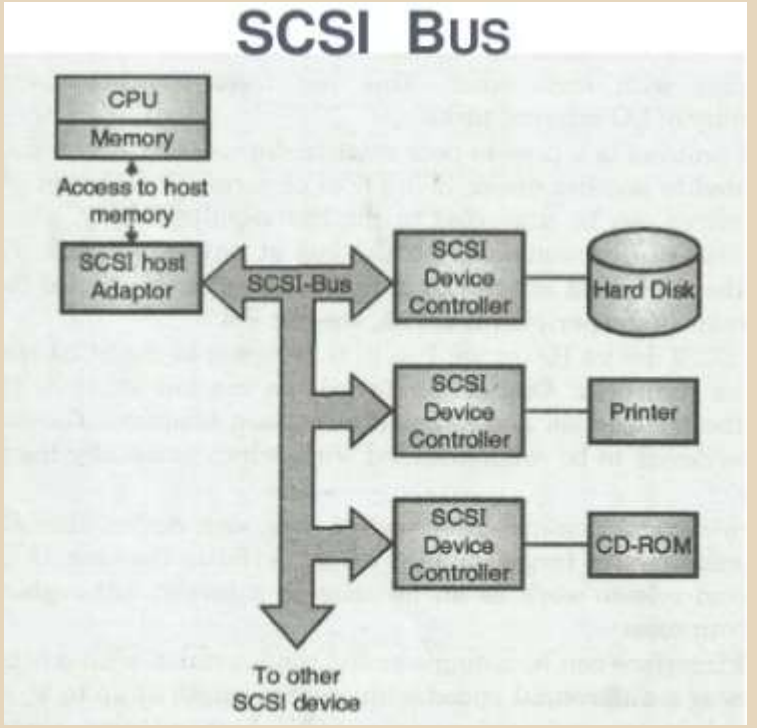
A motherboard with two 32-bit PCI slots and two sizes of PCI Express slots



SCSI Bus

- ✓ **Small Computer System Interface (SCSI)** is a set of standards for physically connecting and transferring data between computers and peripheral devices.
- ✓ The SCSI standards define commands, protocols, electrical, optical and logical interfaces.
- ✓ SCSI is most commonly used for hard disk drives and tape drives, but it can connect a wide range of other devices, including scanners and CD drives, although not all controllers can handle all devices.

SCSI Bus

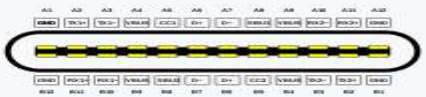


SCSI bus with host adapter and device controller

USB

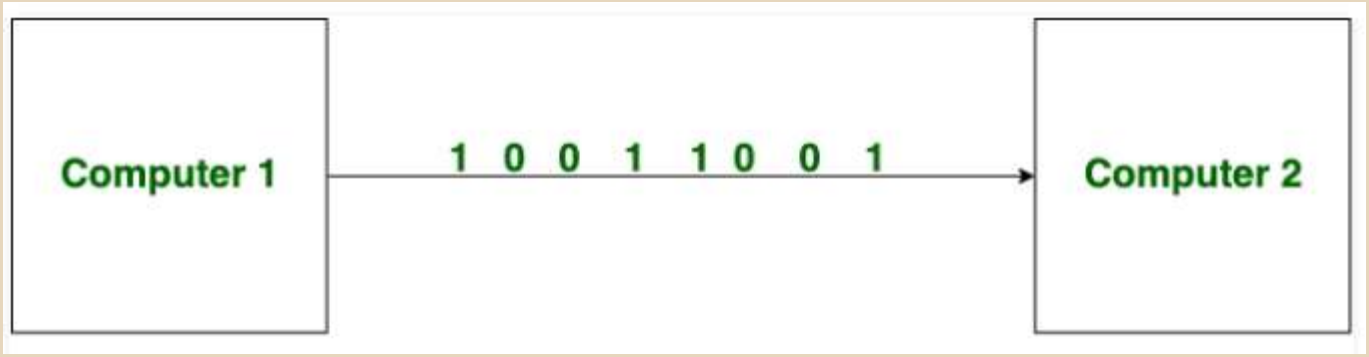
- ✓ **Universal Serial Bus** (**USB**) is an industry standard that establishes specifications for cables and connectors and protocols for connection, communication and power supply (interfacing) between computers, peripherals and other computers.
- ✓ USB was designed to standardize the connection of peripherals to personal computers, both to communicate with and to supply electric power. It has largely replaced interfaces such as serial ports and parallel ports, and has become commonplace on a wide range of devices.
- ✓ Examples of peripherals that are connected via USB include computer keyboards and mice, video cameras, printers, portable media players, disk drives, and network adapters.

USB

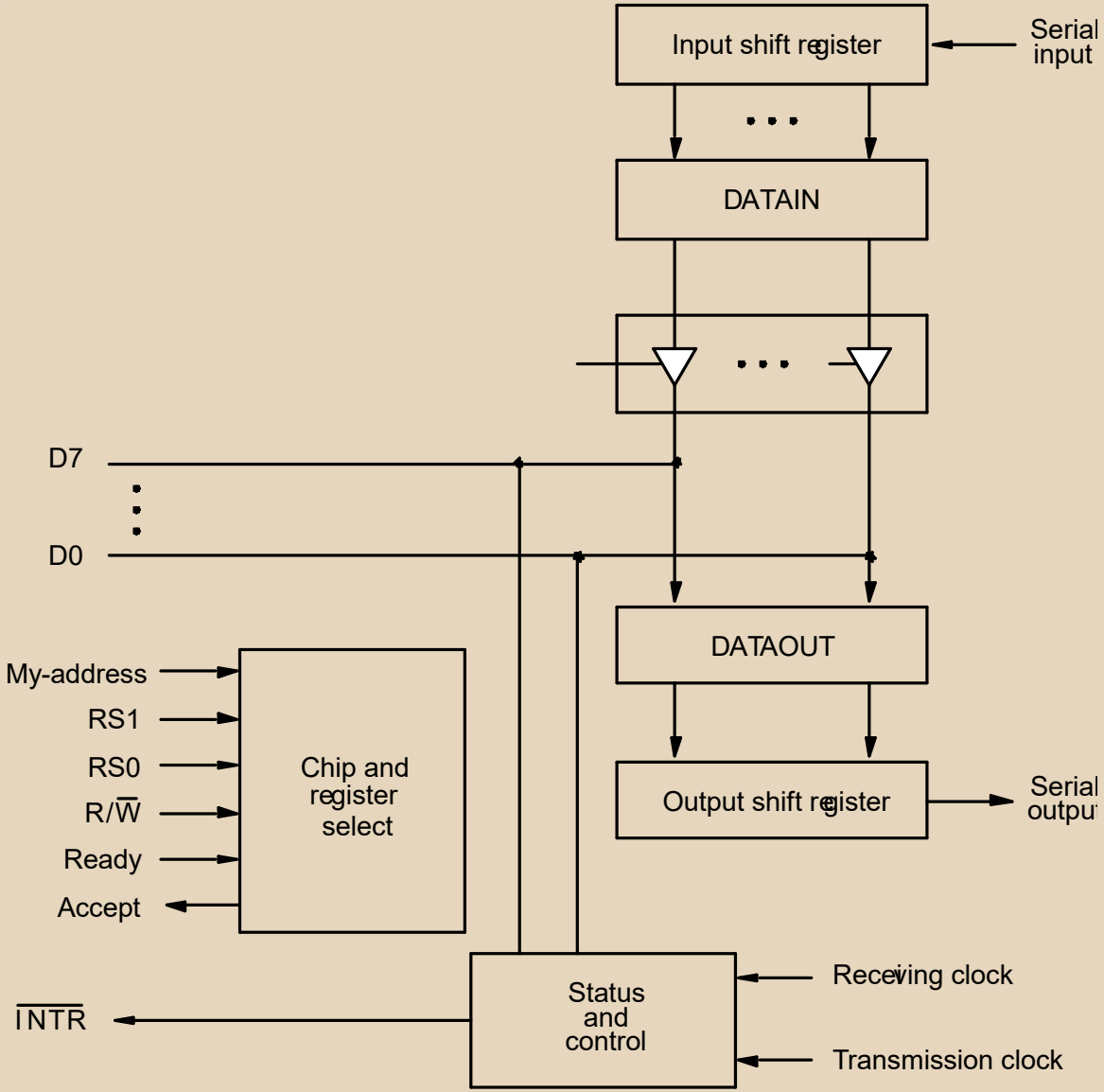
Connectors		USB 1.0 1996	USB 1.1 1998	USB 2.0 2001	USB 2.0 Revised	USB 3.0 2011	USB 3.1 2014	USB 3.2 2017	USB4 2019
Data rate		1.5 Mbit/s (*Low Speed*) 12 Mbit/s (*Full Speed*)	1.5 Mbit/s (*Low Speed*) 12 Mbit/s (*Full Speed*)	1.5 Mbit/s (*Low Speed*) 12 Mbit/s (*Full Speed*) 480 Mbit/s (*High Speed*)		5 Gbit/s (*SuperSpeed*)	10 Gbit/s (*SuperSpeed+*)	20 Gbit/s (*SuperSpeed+*)	40 Gbit/s (*SuperSpeed+ and Thunderbolt 3*)
Standard	A	Type A 				Type A 		Deprecated	
	B	Type B 				Type B 		Deprecated	
	C	N/A		Type C (*enlarged*) 					
Mini	A	N/A	Mini A 			Deprecated			
	B		Mini B 						
	AB	N/A			Mini AB 				
Micro	A	N/A							
	B	N/A			Micro B 	Micro B 		Deprecated	
	AB				Micro AB 	Deprecated			
Connectors		USB 1.0 1996	USB 1.1 1998	USB 2.0 2001	USB 2.0 Revised	USB 3.0 2011	USB 3.1 2014	USB 3.2 2017	USB4 2019

Serial Data Transfer

- ✓ A serial port is used to connect the processor to I/O devices that require transmission of data one bit at a time.
- ✓ In Serial Transmission, data-bit flows from one computer to another computer in bi-direction. In this transmission, one bit flows at one clock pulse. In Serial Transmission, 8 bits are transferred at a time having a start and stop bit.

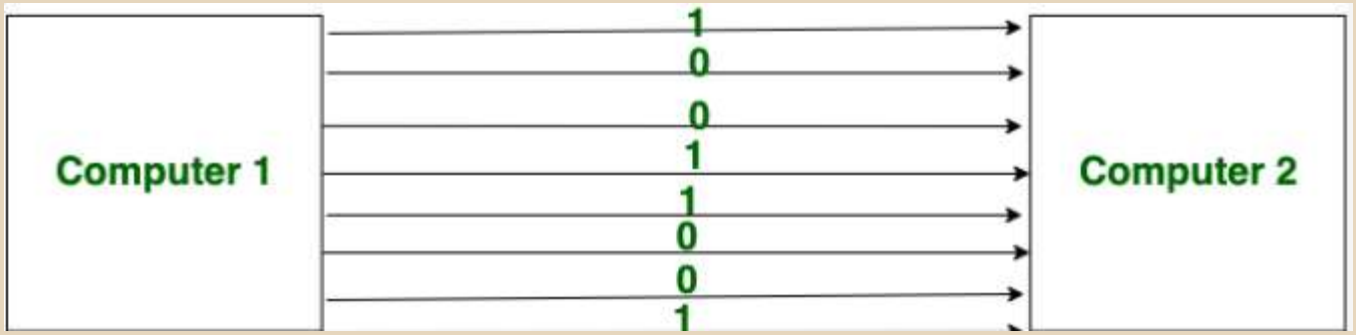


Serial Data Transfer

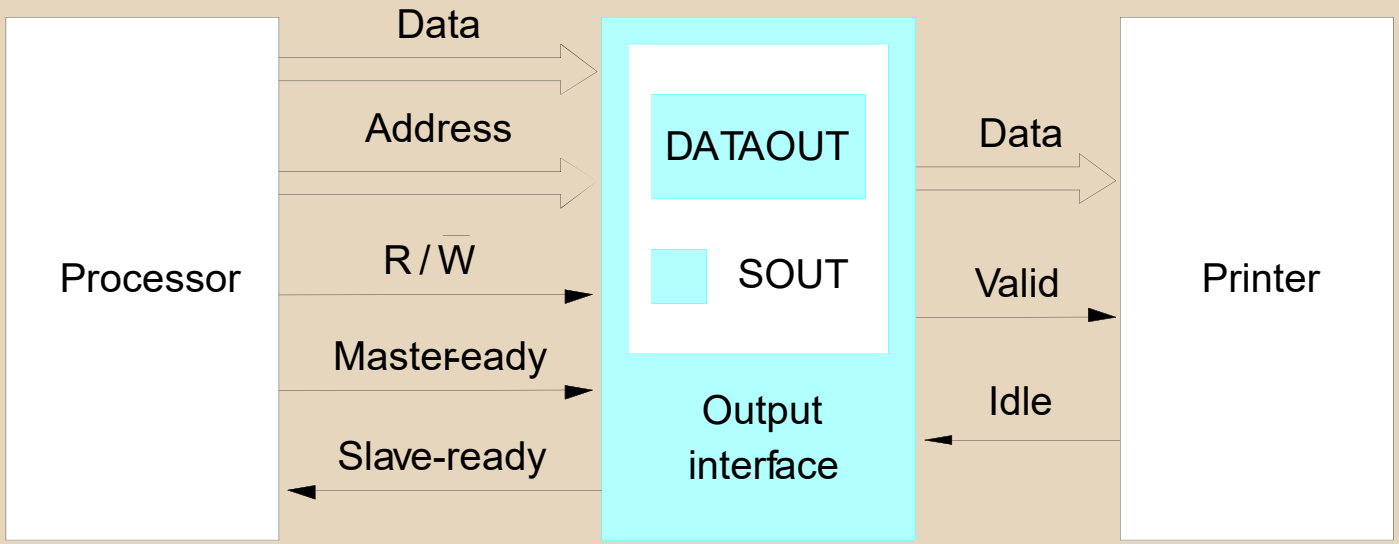


Parallel Data Transfer

- ✓ A parallel port transfers data in the form of a number of bits, typically 8 or 16, simultaneously to or from the device.
- ✓ For faster communications
- ✓ In Parallel Transmission, many bits are flow together simultaneously from one computer to another computer. Parallel Transmission is faster than serial transmission to transmit the bits. Parallel transmission is used for short distance.



Parallel Data Transfer



Printer to processor connection.

Serial Vs Parallel Transmission

S.NO	Serial Transmission	Parallel Transmission
1.	In serial transmission, data(bit) flows in bi-direction.	In Parallel Transmission, data flows in multiple lines.
2.	Serial Transmission is cost-efficient.	Parallel Transmission is not cost-efficient.
3.	In serial transmission, one bit transferred at one clock pulse.	In Parallel Transmission, eight bits transferred at one clock pulse.
4.	Serial Transmission is slow in comparison of Parallel Transmission.	Parallel Transmission is fast in comparison of Serial Transmission.
5.	Generally, Serial Transmission is used for long-distance.	Generally, Parallel Transmission is used for short distance.
6.	The circuit used in Serial Transmission is simple.	The circuit used in Parallel Transmission is relatively complex.

Data Transfer Schemes

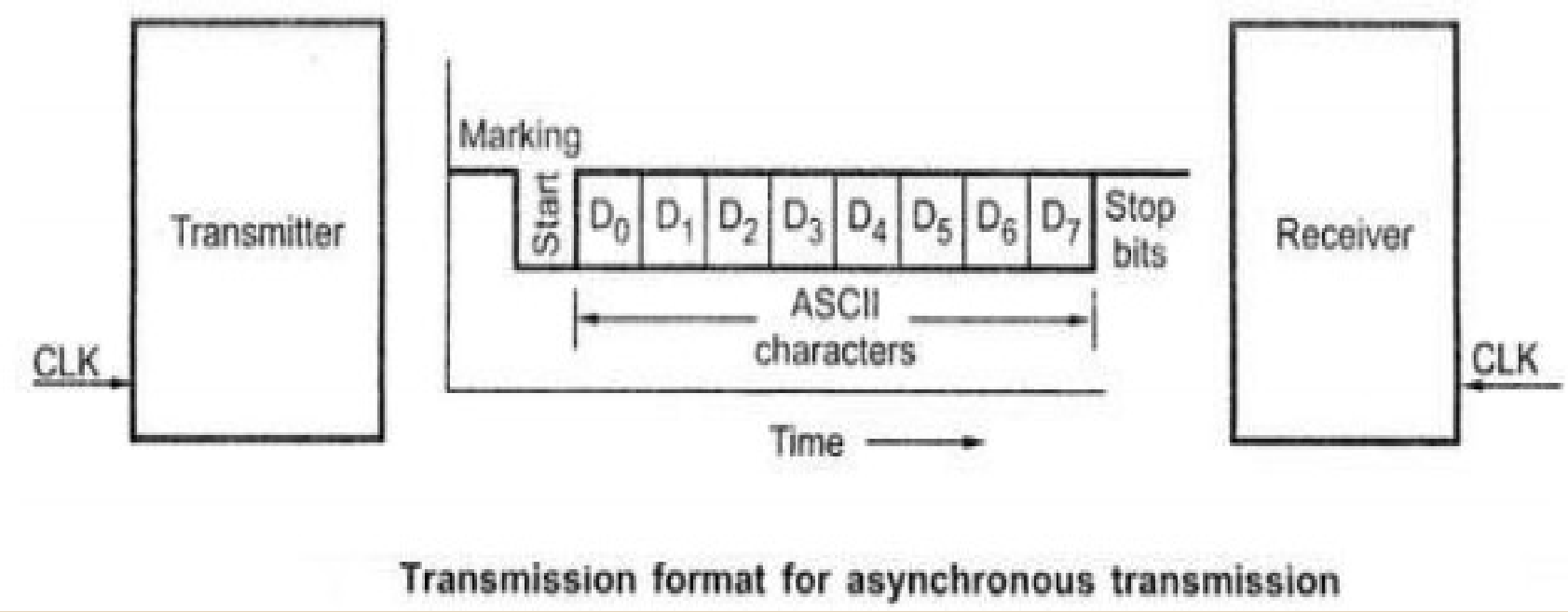
- ✓ Serial communication can be classified on the basis how transmission occurs.
    1. **Simplex**: In simplex, the hardware such that data transfer takes place in only one direction. Ex: Computer to Printer communication
    2. **Half Duplex**: The half duplex transmission allows the data transfer in both direction but not simultaneously. Ex: Walkie talkie
    3. **Full Duplex**: It allows the data transfer in both direction simultaneously. Ex: Telephone lines

- ✓ Data transfer schemes: The data in the serial communication may be sent in two formats:
    1. Asynchronous
    2. Synchronous

Asynchronous Data Transfer

- ✓ Asynchronous formats are character oriented.
- ✓ In this type the bits or character or data word are sent at constant rate, but characters can come at any rate (asynchronously) as long as they do not overlap.
- ✓ When no characters are being sent a line stays high at logic1 called **mark**, logic0 is called **space**.
- ✓ The beginning of a character is indicated by start bit which is always low.
- ✓ This is used to synchronize the transmitter and receiver.
- ✓ After the start bit the data bits are sent with least significant bit first followed by one or more stop bits (active high).
- ✓ The stop bits indicate the end of character.
- ✓ The combination of start bit, character and stop bits is known as frame.
- ✓ The start and stop bits carry no information, but are required because of asynchronous nature of data.

Asynchronous Data Transfer

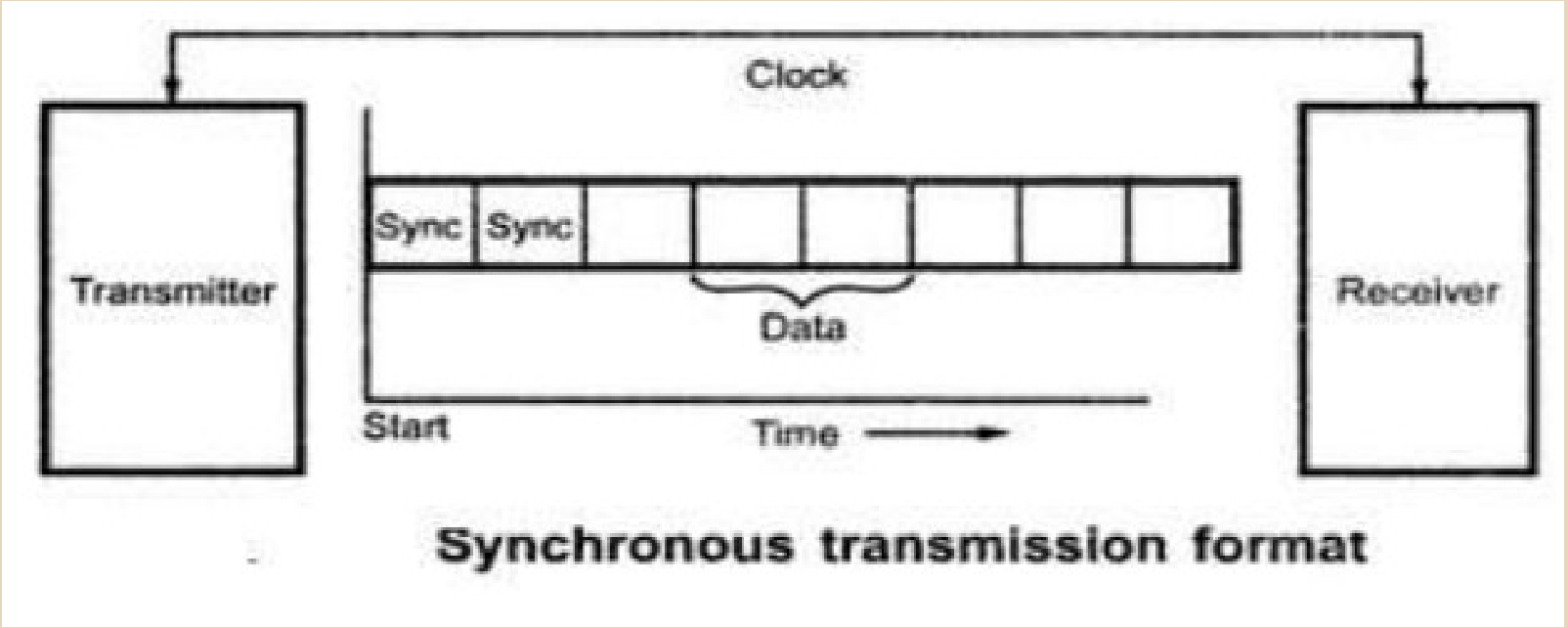


Transmission format for asynchronous transmission

Synchronous Data Transfer

- ✓ The start and stop bits in each frame of asynchronous format represents wasted overhead bytes that reduce overall character rate.
- ✓ These start and stop bits can be eliminated by synchronizing receiver and transmitter.
- ✓ They can be synchronized by having a common clock signal. Such a communication is called synchronous serial communication.
- ✓ In this transmission synchronous bits are inserted instead of start and stop bits
- ✓ The data rate can be expressed as bit/sec or character/sec.
- ✓ The term bit/sec is also called **baud rate**.

Synchronous Data Transfer



Synchronous transmission format

Asynchronous Vs Synchronous

S.No	Asynchronous	Synchronous
1.	Transmitters and receivers are not synchronized by clock.	Transmitters and receivers are synchronized by clock.
2.	Bits of data are transmitted at constant rate.	Data bits are transmitted with synchronization of clock.
3.	Character may arrive at any rate at receiver.	Character is received at constant rate.
4.	Data transfer is character oriented.	Data transfer takes place in blocks
5.	Start and stop bits are required to establish communication of each character.	Start and stop bits are not required to establish communication of each character. Synchronization bits are required to transfer the data block.
6.	Used in low-speed transmission at about speed less than 20 Kbits/sec.	Used in high speed transmissions.

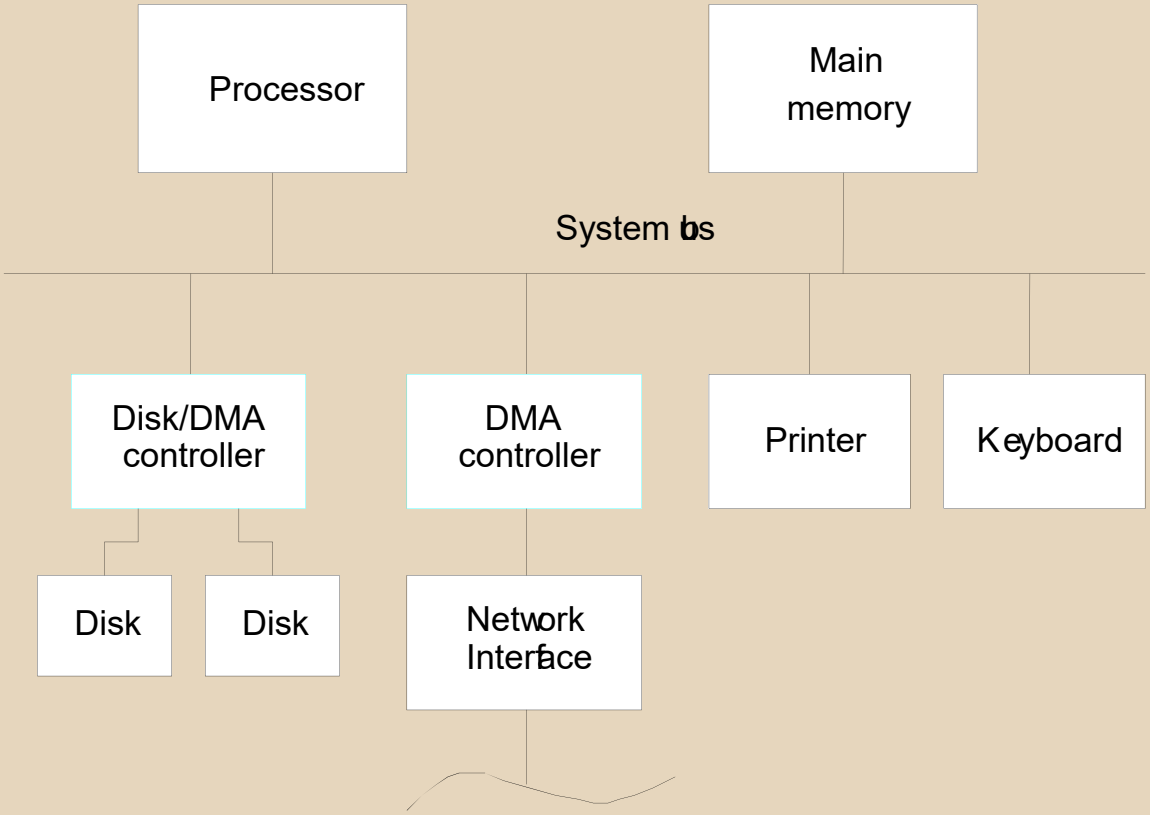
DMA

- ✓ Think about the overhead in both polling and interrupting mechanisms when a large block of data need to be transferred between the processor and the I/O device.
- ✓ A special control unit may be provided to allow transfer of a block of data directly between an external device and the main memory, without continuous intervention by the processor – direct memory access (DMA).
- ✓ The DMA controller provides the memory address and all the bus signals needed for data transfer, increment the memory address for successive words, and keep track of the number of transfers.

DMA Operation

- ✓ Processor sends the starting address, the number of data, and the direction of transfer to DMA controller.
- ✓ Processor suspends the application program requesting DMA, starts DMA transfer, and starts another program.
- ✓ After the DMA transfer is done, DMA controller sends an interrupt signal to the processor.
- ✓ The processor puts the suspended program in the Runnable state.

DMA System



Use of DMA controllers in a computer system.

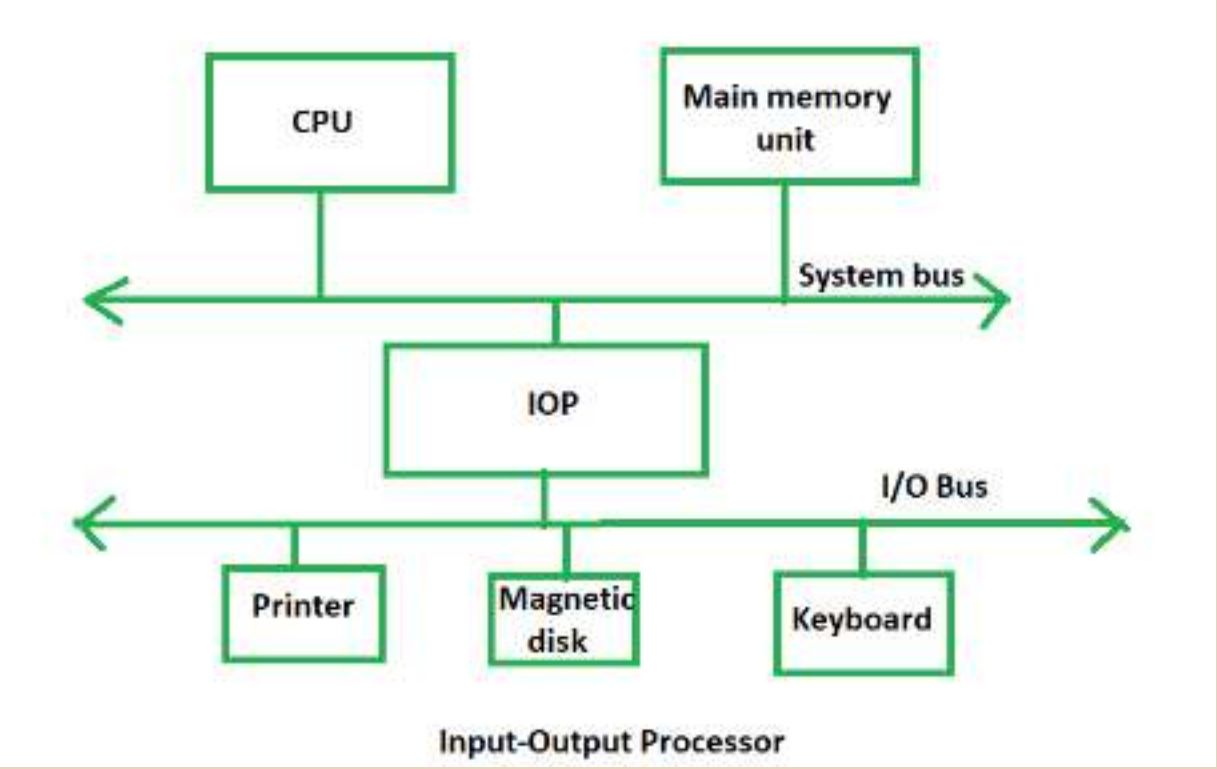
I/O Processor

- ✓ The **DMA mode** of data transfer reduces CPU's overhead in handling I/O operations.
- ✓ It also allows parallelism in CPU and I/O operations. Such parallelism is necessary to avoid wastage of valuable CPU time while handling I/O devices whose speeds are much slower as compared to CPU.
- ✓ The concept of DMA operation can be extended to relieve the CPU further from getting involved with the execution of I/O operations.
- ✓ This gives rises to the development of special purpose processor called **Input-Output Processor (IOP) or IO channel**.

I/O Processor

- ✓ The Input Output Processor (IOP) is just like a CPU that handles the details of I/O operations.
- ✓ It is more equipped with facilities than those are available in typical DMA controller.
- ✓ The IOP can fetch and execute its own instructions that are specifically designed to characterize I/O transfers. In addition to the I/O – related tasks, it can perform other processing tasks like arithmetic, logic, branching and code translation.
- ✓ The main memory unit takes the pivotal role. It communicates with processor by the means of DMA.

I/O Processor



Input-Output Processor

- ✓ The Input Output Processor is a specialized processor which loads and stores data into memory along with the execution of I/O instructions.
- ✓ It acts as an interface between system and devices. It involves a sequence of events to executing I/O operations and then store the results into the memory.

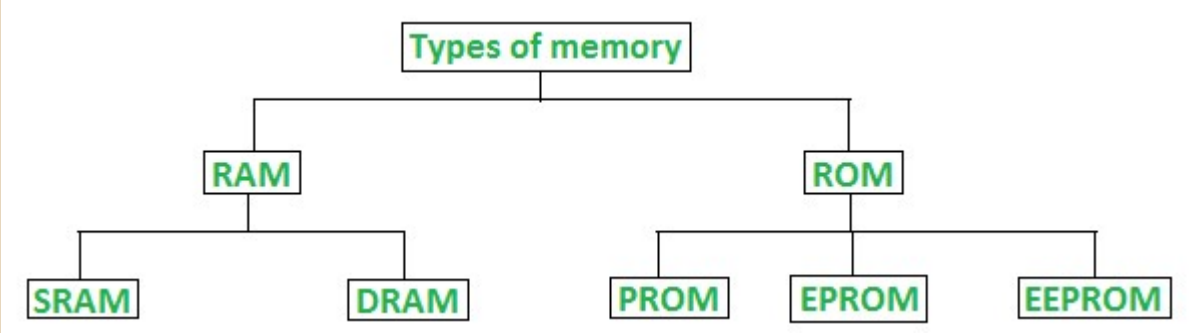
Main memory

JAIN DEEMED-TO-BE UNIVERSITY | SCHOOL OF COMPUTER SCIENCE AND IT

- ✓ A memory is just like a human brain. It is used to store data and instructions.
- ✓ Computer memory is the storage space in the computer, where data is to be processed and instructions required for processing are stored.
- ✓ The memory is divided into large number of small parts called cells.
- ✓ Each location or cell has a unique address, which varies from zero to memory size minus one.
- ✓ For example, if the computer has 64k words, then this memory unit has 64 * 1024 = 65536 memory locations. The address of these locations varies from 0 to 65535.

Main memory

- ✓ Memory is the most essential element of a computing system because without it computer can't perform simple tasks.
- ✓ Computer memory is of two basic type – Primary memory (RAM and ROM) and Secondary memory(hard drive, CD, etc.).
- ✓ Random Access Memory (RAM) is primary-volatile memory and Read Only Memory (ROM) is primary-non-volatile memory.



ASCII Table

Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char
0	0	0		32	20	40	[space]	64	40	100	@	96	60	140	`
1	1	1		33	21	41	!	65	41	101	A	97	61	141	a
2	2	2		34	22	42	"	66	42	102	B	98	62	142	b
3	3	3		35	23	43	#	67	43	103	C	99	63	143	c
4	4	4		36	24	44	$	68	44	104	D	100	64	144	d
5	5	5		37	25	45	%	69	45	105	E	101	65	145	e
6	6	6		38	26	46	&	70	46	106	F	102	66	146	f
7	7	7		39	27	47	'	71	47	107	G	103	67	147	g
8	8	10		40	28	50	(	72	48	110	H	104	68	150	h
9	9	11		41	29	51	)	73	49	111	I	105	69	151	i
10	A	12		42	2A	52	*	74	4A	112	J	106	6A	152	j
11	B	13		43	2B	53	+	75	4B	113	K	107	6B	153	k
12	C	14		44	2C	54	,	76	4C	114	L	108	6C	154	l
13	D	15		45	2D	55	-	77	4D	115	M	109	6D	155	m
14	E	16		46	2E	56	.	78	4E	116	N	110	6E	156	n
15	F	17		47	2F	57	/	79	4F	117	O	111	6F	157	o
16	10	20		48	30	60	0	80	50	120	P	112	70	160	p
17	11	21		49	31	61	1	81	51	121	Q	113	71	161	q
18	12	22		50	32	62	2	82	52	122	R	114	72	162	r
19	13	23		51	33	63	3	83	53	123	S	115	73	163	s
20	14	24		52	34	64	4	84	54	124	T	116	74	164	t
21	15	25		53	35	65	5	85	55	125	U	117	75	165	u
22	16	26		54	36	66	6	86	56	126	V	118	76	166	v
23	17	27		55	37	67	7	87	57	127	W	119	77	167	w
24	18	30		56	38	70	8	88	58	130	X	120	78	170	x
25	19	31		57	39	71	9	89	59	131	Y	121	79	171	y
26	1A	32		58	3A	72	:	90	5A	132	Z	122	7A	172	z
27	1B	33		59	3B	73	;	91	5B	133	[	123	7B	173	{
28	1C	34		60	3C	74	<	92	5C	134	\	124	7C	174	|
29	1D	35		61	3D	75	=	93	5D	135	]	125	7D	175	}
30	1E	36		62	3E	76	>	94	5E	136	^	126	7E	176	~
31	1F	37		63	3F	77	?	95	5F	137	_	127	7F	177	

RAM

- ✓ Random Access Memory (RAM) is also called as read write memory or the main memory or the primary memory.
- ✓ The programs and data that the CPU requires during execution of a program are stored in this memory.
- ✓ It is a volatile memory as the data loses when the power is turned off.
- ✓ RAM is further classified into two types- SRAM (Static Random Access Memory) and DRAM (Dynamic Random Access Memory).

RAM

JGI JAIN DEEMED-TO-BE UNIVERSITY | SCHOOL OF COMPUTER SCIENCE AND IT

DRAM	SRAM
1. Constructed of tiny capacitors that leak electricity.	1.Constructed of circuits similar to D flip-flops.
2.Requires a recharge every few milliseconds to maintain its data.	2.Holds its contents as long as power is available.
3.Inexpensive.	3.Expensive.
4. Slower than SRAM.	4. Faster than DRAM.
5. Can store many bits per chip.	5. Can not store many bits per chip.
6. Uses less power.	6.Uses more power.
7.Generates less heat.	7.Generates more heat.
8. Used for main memory.	8. Used for cache.

ROM

- ✓ Stores crucial information essential to operate the system, like the program essential to boot the computer.
- ✓ It is not volatile.
- ✓ Always retains its data.
- ✓ Used in embedded systems or where the programming needs no change.
- ✓ Used in calculators and peripheral devices.
- ✓ ROM is further classified into 3 types- *PROM*, *EPROM*, and *EEPROM*.

ROM

JAIN DEEMED-TO-BE UNIVERSITY | SCHOOL OF COMPUTER SCIENCE AND IT

- ✓ **Types of Read Only Memory (ROM) –**
- ✓ **PROM (Programmable read-only memory)** – It can be programmed by user. Once programmed, the data and instructions in it cannot be changed.
- ✓ **EPROM (Erasable Programmable read only memory)** – It can be reprogrammed. To erase data from it, expose it to ultra violet light. To reprogram it, erase all the previous data.
- ✓ **EEPROM (Electrically erasable programmable read only memory)** – The data can be erased by applying electric field, no need of ultra violet light. We can erase only portions of the chip.

RAM Vs ROM

JGI JAIN DEEMED-TO-BE UNIVERSITY SCHOOL OF COMPUTER SCIENCE AND IT

RAM	ROM
1. Temporary Storage.	1. Permanent storage.
2. Store data in MBs.	2. Store data in GBs.
3. Volatile.	3. Non-volatile.
4.Used in normal operations.	4. Used for startup process of computer.
5. Writing data is faster.	5. Writing data is slower.

Secondary Memory

JAIN DEEMED-TO-BE UNIVERSITY | SCHOOL OF COMPUTER SCIENCE AND IT

- ✓ This type of memory is also known as external memory or non-volatile.
- ✓ It is slower than the main memory.
- ✓ These are used for storing data/information permanently.
- ✓ CPU directly does not access these memories, instead they are accessed via input-output routines.
- ✓ The contents of secondary memories are first transferred to the main memory, and then the CPU can access it.
- ✓ For example, disk, CD-ROM, DVD, Optical Storage etc.

Secondary Memory

JAIN DEEMED-TO-BE UNIVERSITY | SCHOOL OF COMPUTER SCIENCE AND IT

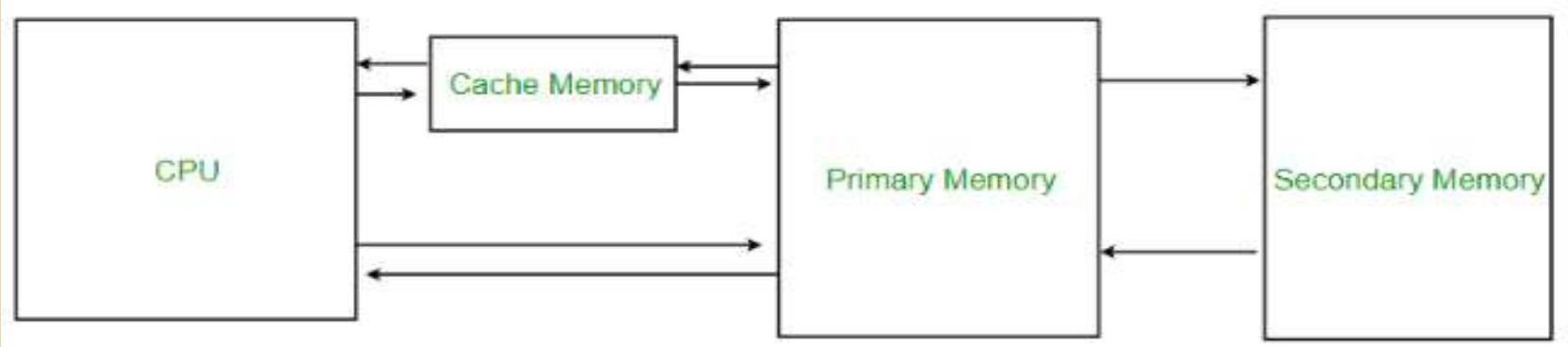
- ✓ Characteristics of Secondary Memory
  - ❖ These are magnetic and optical memories.
  - ❖ It is known as the backup memory.
  - ❖ It is a non-volatile memory.
  - ❖ Data is permanently stored even if power is switched off.
  - ❖ It is used for storage of data in a computer.
  - ❖ Computer may run without the secondary memory.
  - ❖ Slower than primary memories.

Cache Memory

JGI JAIN DEEMED-TO-BE UNIVERSITY SCHOOL OF COMPUTER SCIENCE AND IT

- ✓ Cache memory is a very high speed semiconductor memory which can speed up the CPU.
- ✓ It acts as a buffer between the CPU and the main memory.
- ✓ It is used to hold those parts of data and program which are most frequently used by the CPU.
- ✓ Cache memory is used to reduce the average time to access data from the Main memory.
- ✓ The cache is a smaller and faster memory which stores copies of the data from frequently used main memory locations.

Cache Memory



✓ Levels of memory:

- **Level 1 or Register –**
  It is a type of memory in which data is stored and accepted that are immediately stored in CPU. Most commonly used register is accumulator, Program counter, address register etc.
- **Level 2 or Cache memory –**
  It is the fastest memory which has faster access time where data is temporarily stored for faster access.
- **Level 3 or Main Memory –**
  It is memory on which computer works currently. It is small in size and once power is off data no longer stays in this memory.
- **Level 4 or Secondary Memory –**
  It is external memory which is not as fast as main memory but data stays permanently in this memory.