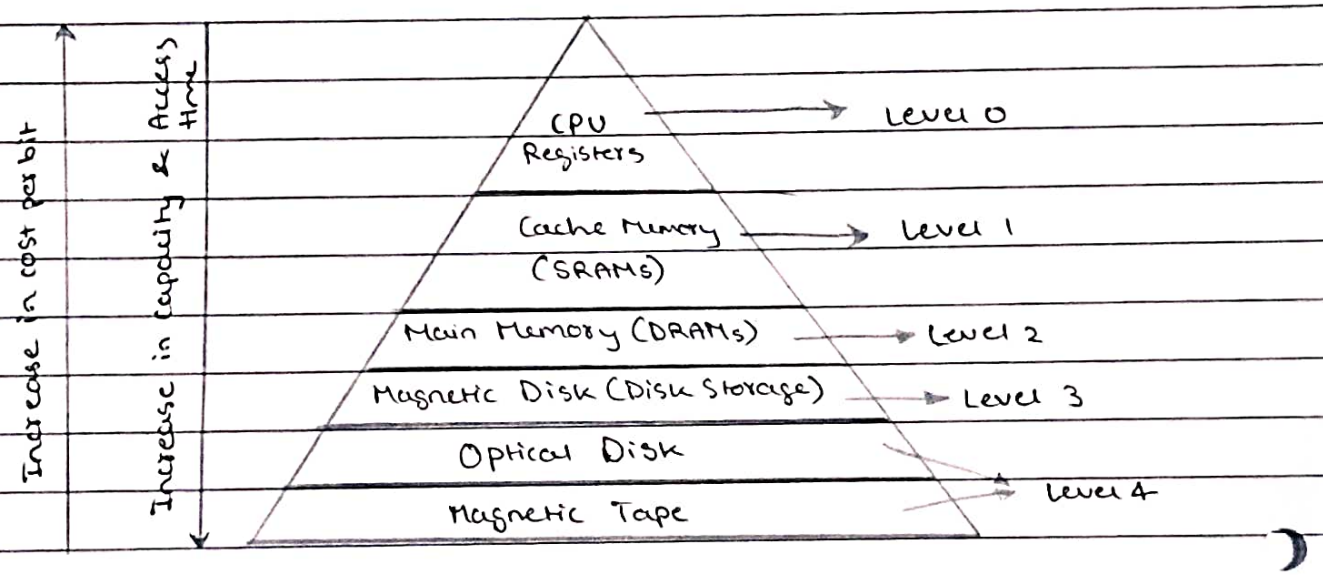
Q1. Compare SRAM and DRAM

SRAM	DRAM
① It stores information as long as the power is supplied.	① It stores information as long as the power is supplied or a few milliseconds when power is switched off
② Transistors are used to store information is SRAM	② Capacitors are used to store data in DRAM
③ Capacitors are not used hence no refreshing is required.	③ To store information for a longer time, contents of the capacitor need to be refreshed periodically.
④ SRAM is faster compared to DRAM	④ DRAM provides slow access speeds
⑤ It does not have a refreshing unit.	⑤ It has a refreshing unit.
⑥ These are expensive	⑥ These are cheaper
⑦ SRAMs are low density devices.	⑦ DRAMs are high-density devices.
⑧ These are used in cache memories	⑧ These are used in main memories
⑨ In this bits are stored in voltage form	⑨ In this bits are stored in form of electric energy
⑩ Consumes less power and generates less heat.	⑩ Uses more power and generates more heat.

Q2 Explain memory hierarchy of a computer

Ans In the Computer System Design, Memory Hierarchy is an enhancement to organize the memory such as that it can minimize the access time. The Memory Hierarchy was developed based on a program behavior known as locality of references. The figure below clearly demonstrates the different levels of memory hierarchy:



This Memory Hierarchy Design is divided into 2 main types:

(1) External Memory or Secondary Memory: Comprising of Magnetic Disk, Optical Disk, Magnetic Tape i.e peripheral storage devices which are accessible by the processor via I/O Module.

(2) Internal Memory or Primary Memory:
Comprising of Main Memory, Cache Memory & CPU registers. This is directly accessible by the processor.

There are typically four levels of memory in a memory hierarchy:

Registers: Registers are small, high-speed memory units located in the CPU. They are used to store the most frequently used data and instructions. Registers have the fastest access time and the smallest storage capacity, typically ranging from 16 to 64 bits.

Cache Memory: Cache memory is a small, fast memory unit located close to the CPU. It stores frequently used data and instructions that have been recently accessed from the main memory. Cache memory is designed to minimize the time it takes to access data by providing the CPU with quick access to frequently used data.

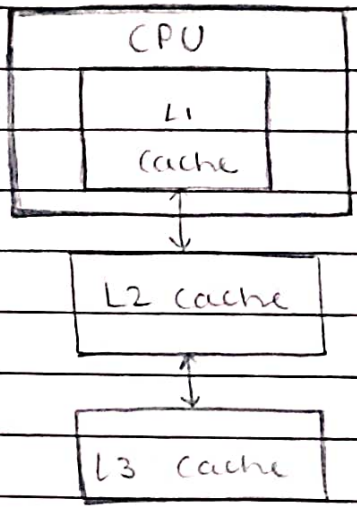
Main Memory: Main memory, also known as RAM (Random Access Memory), is the primary memory of a computer system. It has a larger storage capacity than cache memory, but it is slower. Main memory is used to store data and instructions that are currently in use by the CPU.

Secondary Storage: Secondary storage, such as hard disk drives (HDD) and solid-state drives (SSD), is a non-volatile memory unit that has a larger storage capacity than main memory. It is used to store data and instructions that are not currently in use by the CPU. Secondary storage has the slowest access time and is typically the least expensive type of memory in the memory hierarchy.

Q3 Write short notes on L1, L2, L3 & Cache memory.

Ans (1) Cache is a small but fast memory device that the CPU can access at relatively faster speeds and that holds a subset of the data in the main memory. They store information the CPU is most likely to need next. If the data item is present in the cache, it is termed as a cache hit otherwise it is a cache miss. A typical organization is shown in figure

(2) The same reason we have a memory hierarchy prompts us to have multi-level caches as well. A cache miss, on the other hand, means the CPU has to go scampering off to find the data elsewhere. This is where the L2 cache comes into play- while it's slower, it's also much larger. If data can't be found in the L2 cache, the CPU continues down the chain to L3 and so on.



① L1 caches are designed to be the fastest as they are closet to the CPU and thus it will be accessed typically by the CPU and the access time of L1 cache has a major effect on the clock rate of the CPU. L1 caches are normally much smaller than the other levels of cache but is much bigger than the CPU's registers. L1 cache is normally on chip with processor as shown in the diagram.

② However recent processors are also known to have multiple levels of cache on the processor chip. Since the processor chip needs to be of a certain size, this highly limits the size of on chip cache. Thus external cache levels are also fairly common. L1 is the smallest in size and gives fastest access.

③ L2 on the other hand is relatively slower but is bigger in size giving higher hit rates. L3 is slower as far as the access time is considered (not as slow as the main memory) and even bigger when

the size is considered. This continues for all the cache levels.

④ Some processors use an inclusive cache design (meaning data stored in the L1 cache is also duplicated in the L2 cache) while others have an exclusive cache design (meaning the two caches never share data).

⑤ The access time of the caches also depends on whether the caches are on chip with the processor or are external to the processor.

⑥ If u consider a three level cache memory having L1, L2 and L3, the average access time (t) is given by:

$$t = h_1(1 + (1-h_1)h_2(2 + (1-h_1)(1-h_2)(3 + (1-h_1)(1-h_2)(1-h_3)M$$

h_1 = Hit rate in L1

h_2 = Hit rate in L2

h_3 = Hit rate in L3

C_1 = Access Time of L1

C_2 = Access Time of L2

C_3 = Access Time of L3

M = Access Time of Main memory.

The number of misses in the L2 and L3 cache given be the terms $(1-h_1)(1-h_2)$ and $(1-h_1)(1-h_2)(1-h_3)$ should be as low as possible. This is the reason why L2 and L3 are larger to get higher hit rates and lower miss rates, resulting in very few main memory accesses.

Q4 Write short note on associative memory.

Ans Associative memory is also known as content addressable memory (CAM) or associative storage or associative array. It is a special type of memory that is optimized for performing searches through data, as opposed to providing a simple direct access to the data based on the address.
It can store the set of patterns as memories when the associative memory is being presented with a key pattern, it responds by producing one of the stored pattern which closely resembles or relates to the key pattern.
It can be viewed as data correlation here. Input data is correlated with that of stored data in the CAM.
It forms of two types:
(1) Auto associative memory network.
(2) Hetero associative memory network.

Associative memory of conventional semiconductor memory (usually RAM) with added comparison circuitry that enables a search operation to complete in a single clock cycle. It is a hardware search engine, a special type of computer memory used in certain very high searching applications.

* Applications of associative memory:
(1) It can be only used in memory allocation format.

(2) It is widely used in the database management systems, etc.

★ Advantages of Associative memory:

(1) It is used where search time needs to be less or short.
(2) It is suitable for parallel searches.
(3) It is often used to speedup databases.
(4) It is often in page tables used by the virtual memory and used in neural networks.

★ Disadvantages of Associative memory:

(1) It is more expensive than RAM.
(2) Each cell must have storage capability and logical circuits for matching its content with external argument.