**Principle of Locality:** The principle of locality states that programs access a relatively small portion of their address space at any instant of time.

**Temporal Locality**: The principle stating that if a data location is referenced then it will tend to be referenced again soon.

**Spatial locality**: The locality principle stating that if a data location is referenced, data locations with nearby addresses will tend to be referenced soon.

**Memory Hierarchy**: A structure that uses multiple levels of memories; as the distance from the processor increases, the size of the memories and the access time both increases. The faster memories are more expensive per bit than the slower memories and thus are smaller.

**Block (or Line)**: The minimum unit of information that can be either present or not present in a cache.

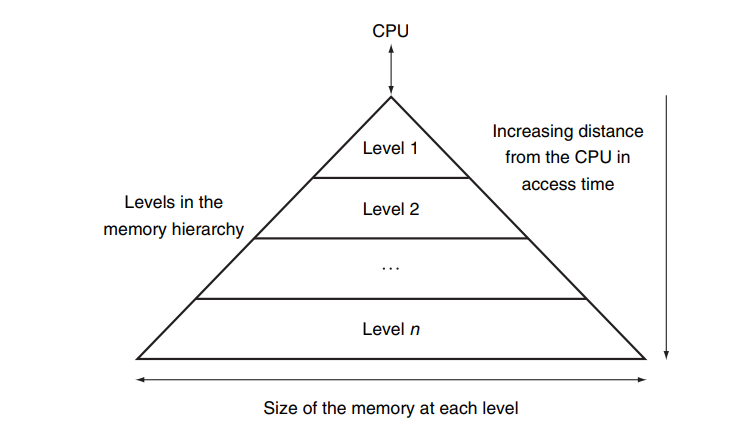
**Hit Rate**: The fraction of memory accesses found in a level of the memory hierarchy.

**Miss Rate**: The fraction of memory accesses not found in a level of the memory hierarchy.

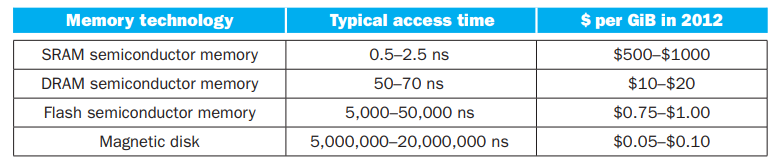
*Hit Rate + Miss Rate = 1*

**Hit Time**: The time required to access a level of the memory hierarchy, including the time needed to determine whether the access is a hit or a miss.

**Miss Penalty**: The time required to fetch a block into a level of the memory hierarchy from the lower level, including the time to access the block, transmit it from one level to the other, insert it in the level that experienced the miss, and then pass the block to the requestor.



**Memory Technologies**



**Track**: One of thousands of concentric circles that makes up the surface of a magnetic disk.

**Sector**: One of the segments that make up a track on a magnetic disk. A sector is the smallest amount of information that is read or written on a disk.

**Cylinder**: The term cylinder is used to refer to all the tracks under the heads at a given point on all surfaces.

**Seek**: The process of positioning a read / write head over the proper track on a disk.

**Seek Time**: The time to move the head to the desired track.

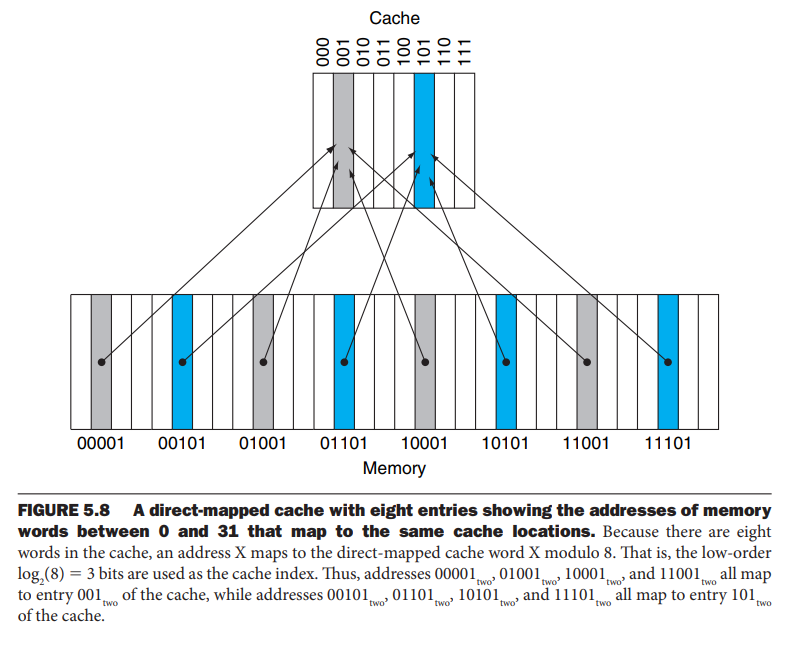
**Rotational latency**: Also called rotational delay. The time required for the desired sector of a disk to rotate under the read / write head. Usually assumed to be half the rotation time.

**Transfer time**: The time to transfer a block of bits. The transfer time is a function of the sector size, the rotation speed, and the recording density of a track.

**Cache**: A safe place for hiding or storing things.

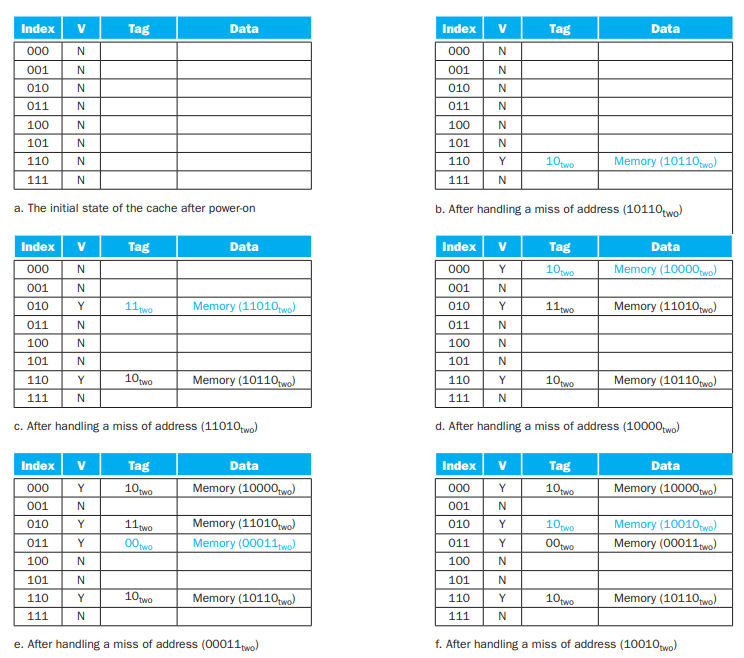
**Direct-mapped cache**: A cache structure in which each memory location is mapped to exactly one location in the cache. Almost all direct-mapped caches use this mapping to find a block: (Block address) modulo (Number of blocks in the cache).

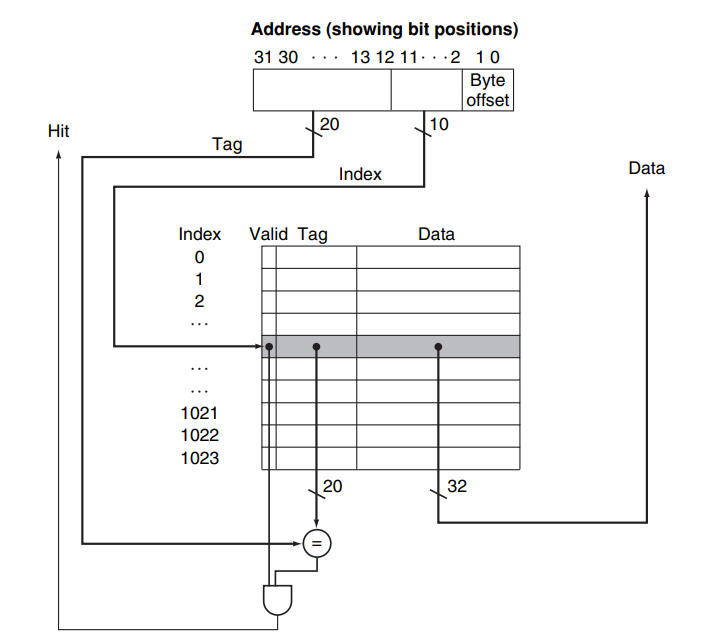
**Tag**: A field in a table used for a memory hierarchy that contains the address information required to identify whether the associated block in the hierarchy corresponds to a requested word.





**Valid Bit**: A field in the tables of a memory hierarchy that indicates that the associated block in the hierarchy contains valid data.





**Cache Miss**: A request for data from the cache that cannot be filled because the data is not present in the cache.

**Write-Through**: A scheme in which writes always update both the cache and the next lower level of the memory hierarchy, ensuring that data is always consistent between the two.

**Write Buffer**: A queue that holds data while the data is waiting to be written to memory.

**Write-Back**: A scheme that handles writes by updating values only to the block in the cache, then writing the modified block to the lower level of the hierarchy when the block is replaced.

**Split Cache**: A scheme in which a level of the memory hierarchy is composed of two independent caches that operate in parallel with each other, with one handling instructions and one handling data.

**Error Detection Code**: A code that enables the detection of an error in data, but not the precise location and, hence, correction of the error.