CACHE MAPPING TECHNIQUES

**Memory mapping** is the (complex) process that associates an address value (usually a 32 or 64 bits number) to some existing physical location in the hardware.This location can be in RAM, in a cache of some level, or even on the hard disk! During program execution, data can move from one location to another, and possibly be duplicated.Mapping Function

The correspondence between the main memory blocks and those in thecache is specified by a mapping function.

The different Cache mapping techniques are as follows:-

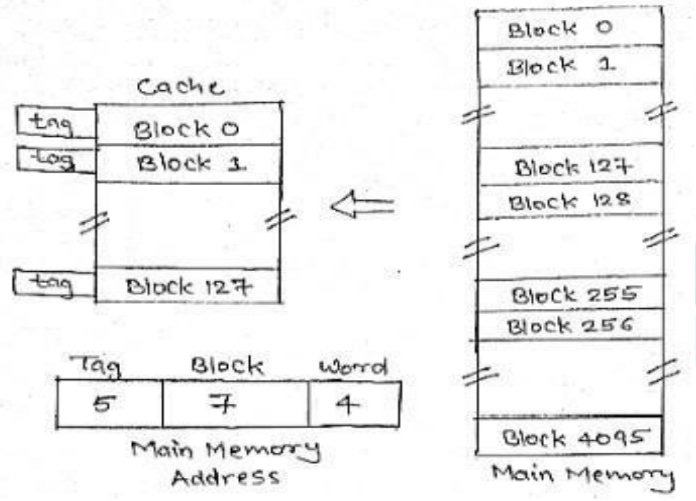
**1)** Direct Mapping

**2) Associative Mapping**

**3)** Set Associative Mapping

Consider a cache consisting of 128 blocks of 16 words each, for total of 2048(2K) works and assume that the main memory is addressable by 16 bit address. Main memory is 64K which will be viewed as 4K blocks of 16 works each.

**(1)** Direct Mapping:-




- The simplest way to determine cache locations in which store Memory blocks isdirect Mapping technique.
- In this block J of the main memory maps on to block J modulo 128 of thecache. Thus main memory blocks 0,128,256,….is loaded into cache is stored at block 0. Block

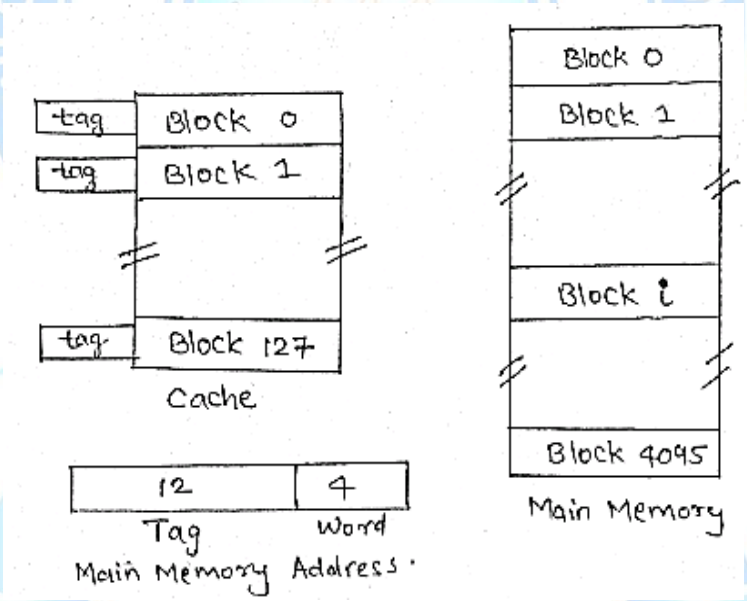
1,129,257,….are stored at block 1 and so on.

- Placement of a block in the cache is determined from memory address. Memoryaddress is divided into 3 fields, the lower 4-bits selects one of the 16 words in a block.
- When new block enters the cache, the 7-bit cache block field determines the cache positions in which this block must be stored.
- The higher order 5- in cache. They identify which of the 32 blocks that are mapped into this cache bits of the memory address of the block are stored in 5 tag bits associated with its location position are currently resident in the cache.

**Advantages:** It is easy to implement.

**Drawbacks:** Since more than one memory block is mapped onto a given cacheblock position, contention may arise for that position even when the cache is not full.Contention is resolved by allowing the new block to overwrite the currently residentblock. This method is not very flexible.

**(2)** Fully Associative Mapping:-

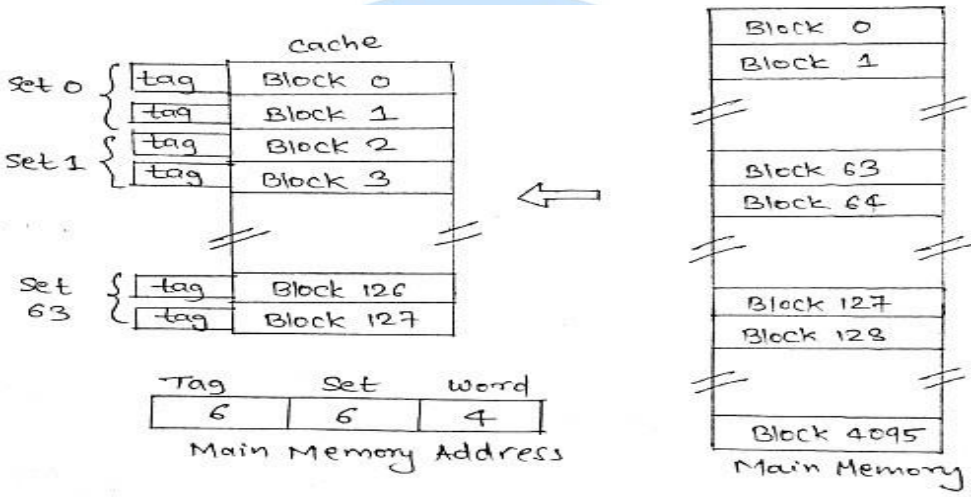



- This is more flexible mapping method, in which main memory block can be placed into any cache block position.
- In this, 12 tag bits are required to identify a memory block when it is resident inthe cache.
- The tag bits of an address received from the processor are compared to the tag bits of each block of the cache to see, if the desired block is present. This is known as Associative Mapping technique.

- Cost of an associated mapped cache is higher than the cost of direct-mapped because of the need to search all 128 tag patterns to determine whether a blockis in cache.This is known as associative search.

**(3)** Set-Associated Mapping:-

- It is the combination of direct and associative mapping technique.
- Cache blocks are grouped into sets and mapping allow block of main memory reside into any block of a specific set. Hence contention problem of directmapping is




eased, at the same time; hardware cost is reduced by decreasing thesize of associative search.

- For a cache with two blocks per set. In this case, memory block 0, 64, 128,…..,4032 map into cache set 0 and they can occupy any two block within this set.
- Having 64 sets means that the 6 bit set field of the address determines which setof the cache might contain the desired block. The tag bits of address must be associatively compared to the tags of the two blocks of the set to check if desiredblock is present. This is known as two way associative search.

Advantages:

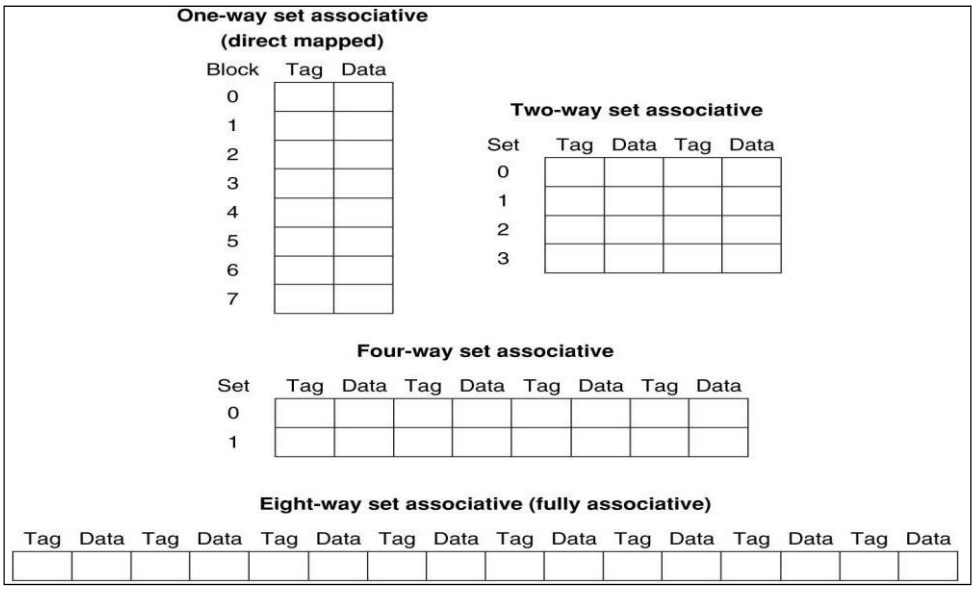
The contention problem of the direct-mapping is eased by having a few choices for block placement. At the same time, the hardware cost is reduced by decreasing the size of the associative search.

M - Way Set Associativity:

We can also think of all block placement strategies as a variation on set

associativity. The following figure shows the possible associativity structures foraneight-block cache.

- A direct-mapped cache is simply a one-way set-associative cache: each cache entry




  holds one block and each set has one element.

- A fully-associative cache with m entries is simply an m-way set-associative cache: it has one set with m blocks, and an entry can reside in any block within that set.

The advantage of increasing the degree of associativity is that it usually decreases the miss rate. The main disadvantage is the potential increase in the hit time.