Cache mapping techniques :-

Blocks are loaded from main memory to cache memory. Cache mapping decides which block of main memory comes into which block of cache memory:

1) Associative Mapping also called fully associative mapping.
2) Direct Mapping also called One-way set associative mapping
3) set Associative mapping also called two-way set associative Mapping.

→ Mapping techniques trying to balance between Hit Ratio, search-time and Tag ~~time~~ size.

→ Each cache block has a Tag indicating which block of main memory is mapped into that block.

→ A collection of such tags is called the cache directory (similar to a page table)

→ ~~Cash Bl~~ Cache Block (cache lines).

→

Associative Mapping: Any block of main memory can be mapped at any available block of cache memory. There are no rules restricting the mapping at all. This means the full cache is available for mapping hence the name Full Associative.

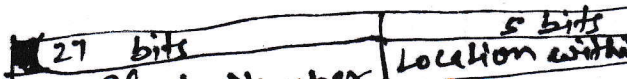
Consider Pentium Processor Cache:

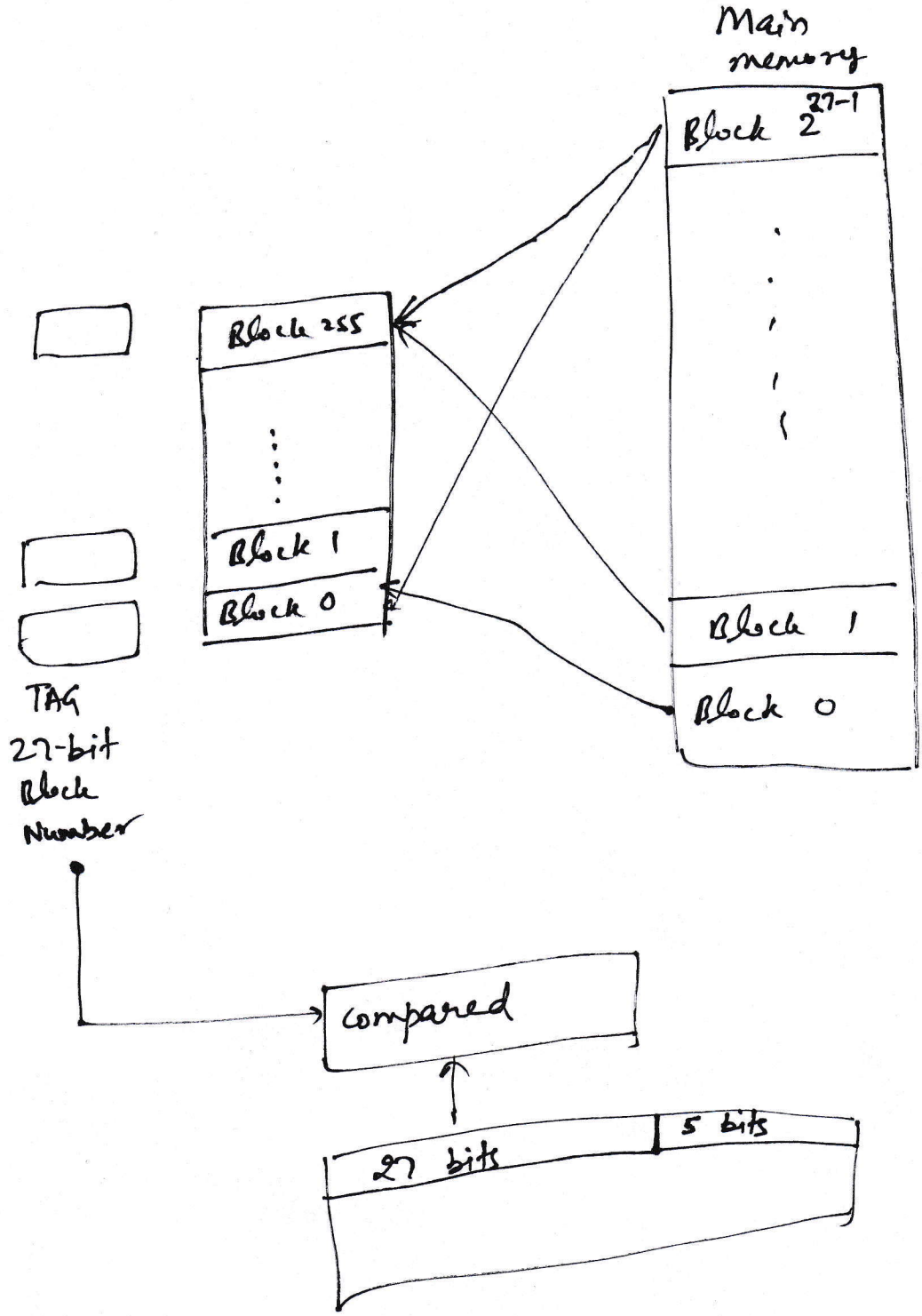
size of Main memory:	$4GB = 2^{32}$
size of cache memory:	$8kB = 2^{13}$
size of cache block (Line):	32 bytes (words) $= 2^5$
No. of blocks in main mem:	size of mm (2^{32}) / size of block $(2^5) = 2^{27}$
No. of blocks in cache Mem:	size of cm (2^{13}) / size of block $(2^5) = 256$
Main Mem address:	32 bits (because mm is of $4GB = 2^{32}$)

Tag size: A block of cache memory can contain any block of main memory out of possible 2^{27} blocks. Hence, the Tag next to every block in Cache Memory must be 27 bits.

Searches: A block of mm can be mapped into any block of cm out of 256 blocks, Hence we need to do 256 searches in cm.

Method of Searching:- The processor issues a 32-bit main memory address. It can be divided as:




Main memory
Block 2^{27-1}
Block 255
Block 1
Block 0
Block 1
Block 0
TAG
27-bit
Block
Number
compared
27 bits
5 bits

Direct Mapping: Any block of main memory can only be mapped at one block of cache memory. Since there is only one way of mapping, it is called one-way set associative Mapping.

→ Entire Cache is treated as one set.

→ The mm is divided into sets which are further subdivided into Blocks.

→ A block of mm (of any set), can only be mapped into the same Block No. in Cache mem.

→ It means, Block 0 of main memory (of any set), can only be mapped into Block 0 of CM.

→ or we can say, Block 0 of CM can only contain Block 0 of main memory but of any set.

Size of mm	$4GB = 2^{32}$
Size of CM	$8kB = 2^{13}$ (It is treated as one set)
Hence, Size of set:	$8kB = 2^{13}$
Size of Cache Block (line):	32 bytes $= 2^5$
No. of blocks in a set:	$2^{13}/2^5 = 2^8 = 256$
No. of sets in mm	$2^{32}/2^{13} = 2^{19}$
No of Sets in CM	1
Main memory address	32 bits

DM

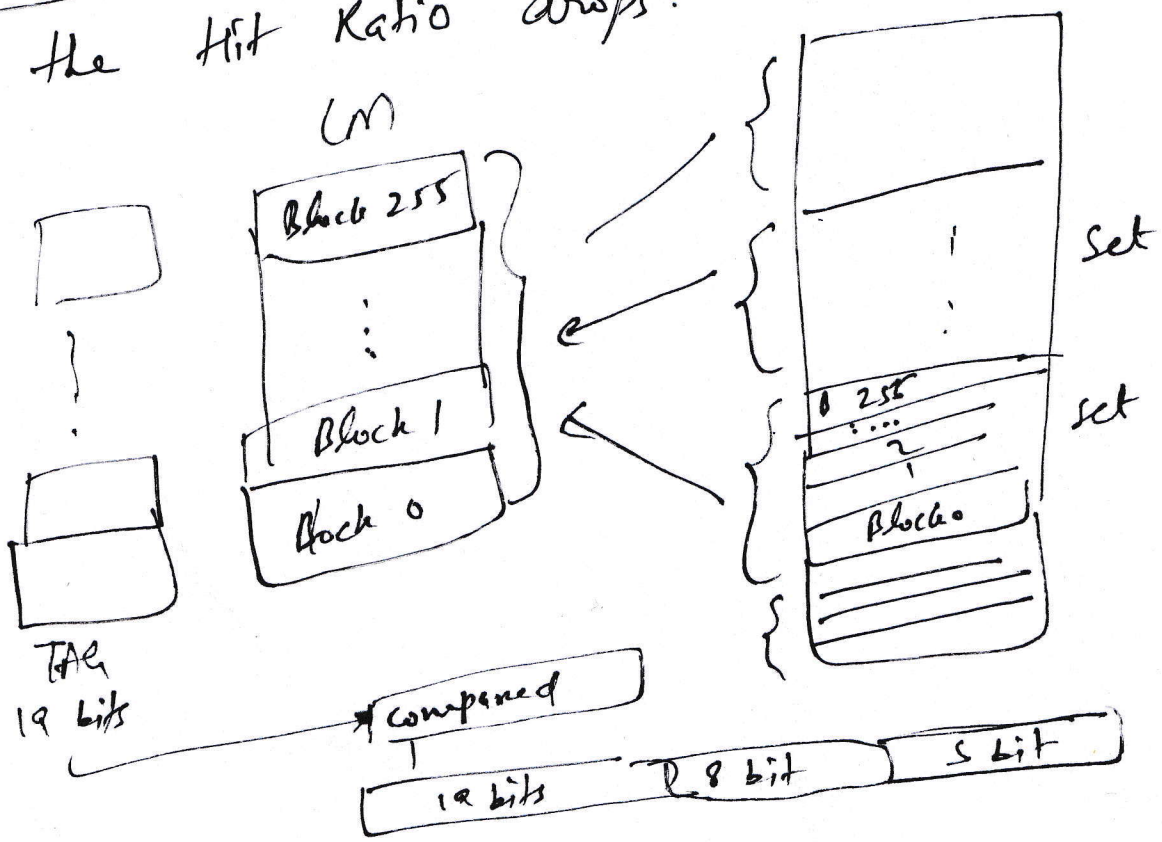
Tag size :- Since, Block 0 of CM can only contain Block 0 of main memory but of any set, the TAG has to only indicate the Set No. if MM. As main memory has 2^{19} sets, the TAG size is 19 bits

Main Memory address	19 bits	8 bit	5 bit
	Set No.	Block No.	Location within Block

Example | 5 : 0 : 6 |

Adv : In one search, we know if it is a Hit or miss. Tag size = 19 bits

Drawback : Since the method is very rigid, the Hit Ratio drops.



Set Associative Mapping : A block of main memory can only be mapped into the same corresponding block No. of cache memory, in any of the two sets.