



Computer Organization and Architecture

Cache memory organization part-1

ABOUT ME : MURALIKRISHNA BUKKASAMUDRAM

- MTech with 20 years of Experience in Teaching GATE and Engineering colleges
- IIT NPTEL Course topper in Theory of computation with 96 %
- IGIP Certified (Certification on International Engineering educator)
- GATE Qualified
- Trained more than 50 Thousand students across the country
- Area of Expertise : TOC,OS,COA,CN,DLD



Cache memory organization part-1

15-Address bits
 12-Data bits

Cache Mapping Techniques

 → phy. Address of Main Memory.

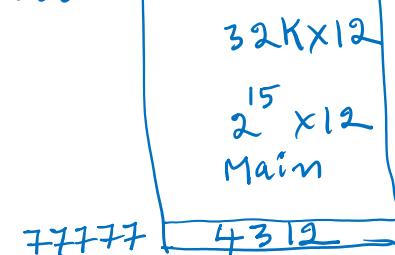
Binary Coded Octal (0 to 7) Main Memory

000 to 07

to 111

000000

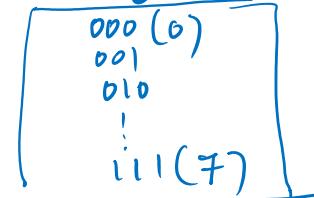
1412 ✓ 00000000000000



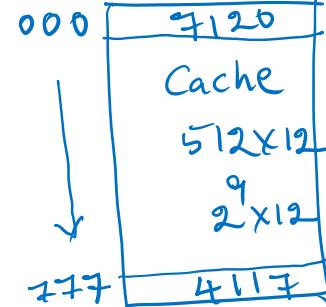
77777

4312

Binary Coded Octal



Cache Memory



 → 15-bits of phy. Mem. Address to the 9-bit Cache Address

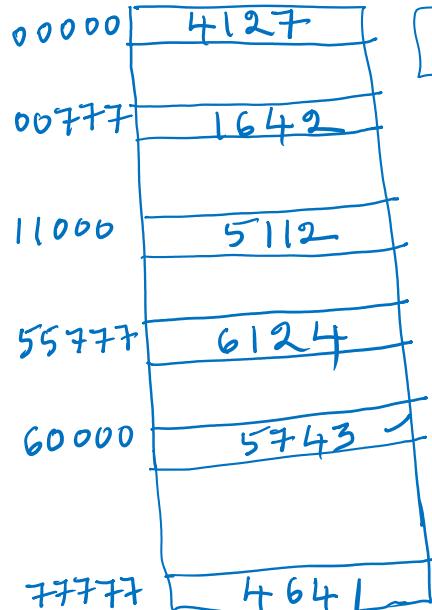
Cache memory organization part-1

Cache Mapping Techniques

- (1) Associative Mapping
- (2) Direct Mapping
- (3) Set-Associative Mapping

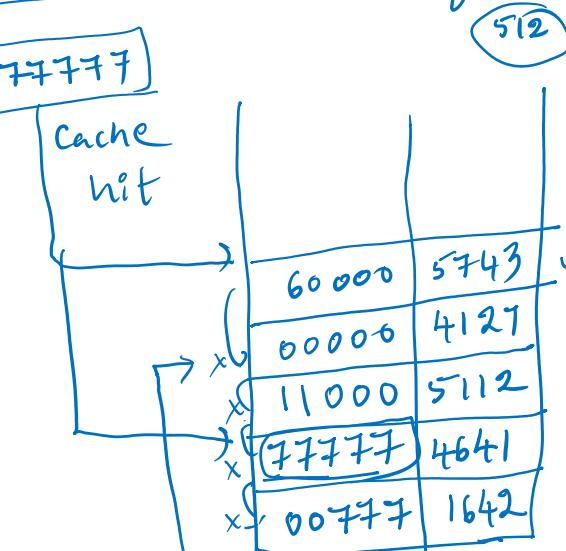
15×512

bits



Main
Memory

Associative Mapping



Cache
Memory

60000

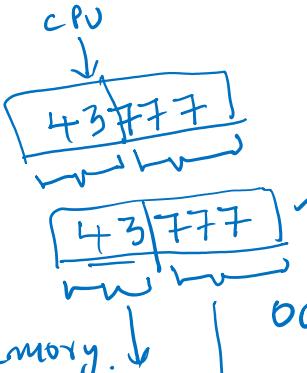
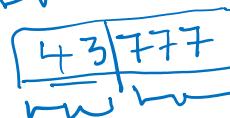
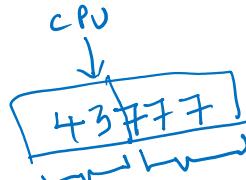
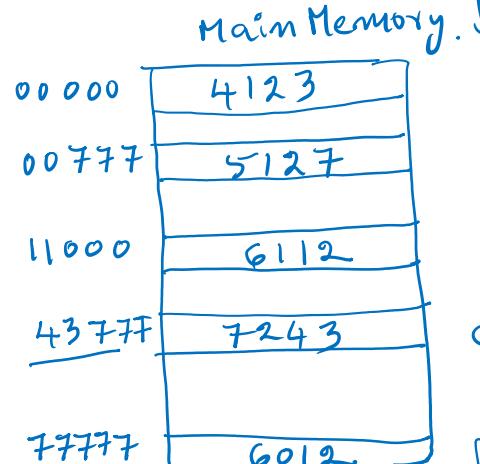
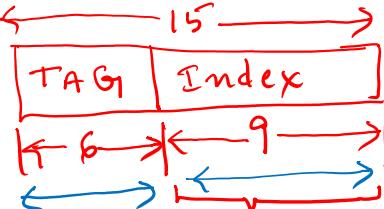
Cache Miss

Cache memory organization part-1

CPU
15-bits
Main
 $32K \times 12$
 $2^{15} \times 12$

Cache
 512×12
 $2^9 \times 12$

Direct Mapping



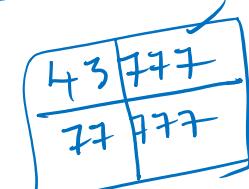
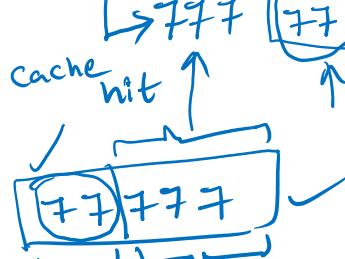
000

Cache Memory

TAG	DATA
11	6112

11000

Cache Miss



512×6 bits

Extra Memory

Cache memory organization part-1

$$32 \text{ KBytes} = 2^{15}$$

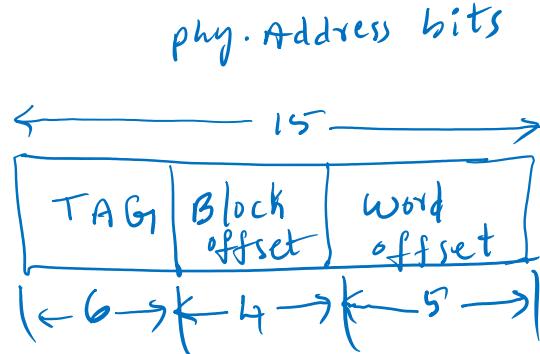
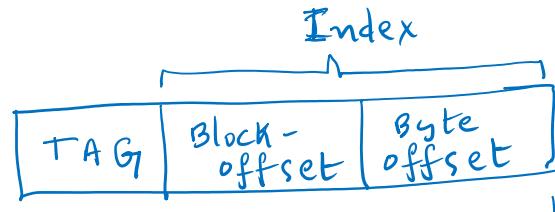
$$512 \text{ Bytes} = 2^9$$

$$\text{Block Size} = 32 \text{ words}$$

Extra Memory required in Cache Controller to store TAGS

Cache memory can also be divided into blocks.

$$\text{No. of Blocks in Cache} = \frac{\text{Cache Size}}{\text{Block Size}}$$



$$\begin{aligned} \text{No. of Blocks} &= \frac{512}{32} \\ &= 2^9 / 2^5 = 2^4 \text{ blocks} \end{aligned}$$

$$\frac{2^5 \text{ words}}{5 - \text{bits}}$$

$$\begin{aligned} 6 \text{ bits} &\times 2^4 \\ 6 \times 16 \text{ bits} &= \underline{96 \text{ bits}} \end{aligned}$$

Cache memory organization part-1

placement of Main Memory blocks in Cache in direct mapping



Initially Cache will be empty.

$$\text{Cache Block No} = \left(\frac{\text{Main Memory Block Number}}{\text{Block Number}} \right) \bmod \left(\frac{\text{No. of Blocks in the Cache}}{\text{Cache Size}} \right)$$

Example :- $\frac{\text{CPU Address}}{4, 1, 5, 9, 8, 15, 16, 12, 13, 3, 5, 4, 12, 3, 9}$

Cache has 8 blocks.

$$4 \bmod 8 = 4 \quad 15 \bmod 8$$

$$1 \bmod 8 = 1$$

$$5 \bmod 8 = 5$$

$$9 \bmod 8 = 1$$

$$8 \bmod 8 = 0$$

0	8	16
1	X	9
2		
3	3	
4	4	+2
5	5	+3
6		
7	15	

Cache memory organization part-1

32

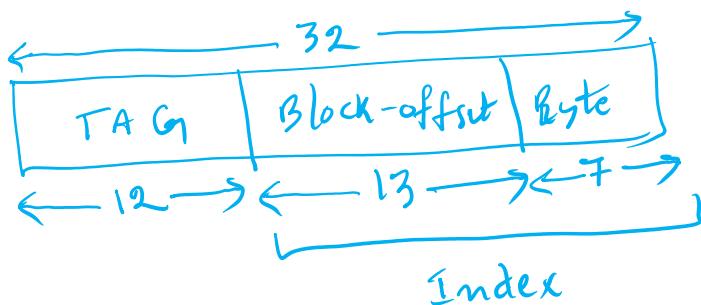


A Direct mapped cache has a size of 1M bytes and divided into blocks of 128 Bytes each. If CPU generates 32-bits addresses, then what is the size of extra memory required in the Cache Controller for storing the TAG info?

Sol

$$\text{No. of Blocks in Cache} = \frac{1\text{M bytes}}{128\text{ bytes}} = \frac{2^{20}}{2^7}$$

$$= \underline{\underline{2^3 \text{ blocks}}}$$



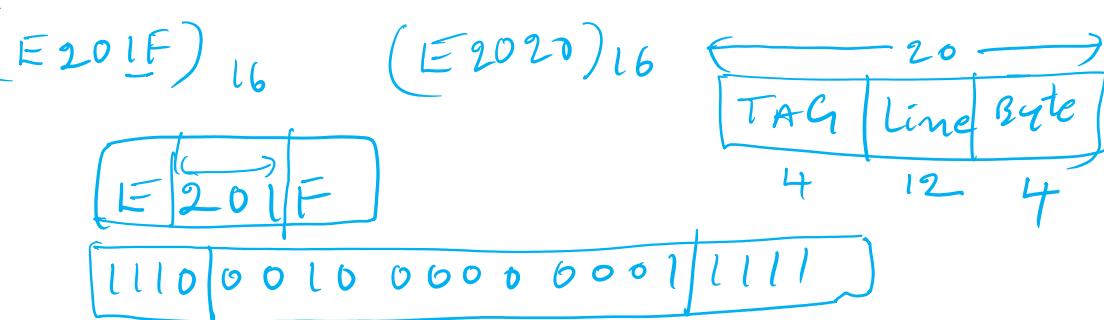
Cache memory organization part-1

$\frac{20}{2} \text{ Bytes}$
 20 bits, 16 Bytes.
 2¹² Lines

→ Consider a machine with byte addressable main memory of 2^{20} bytes, block size of 16 bytes and a direct mapped cache having 2^{12} cache lines. Let the address of two consecutive bytes in main memory be $(E201F)_{16}$ and $(E2020)_{16}$. What are the tag and cache line address (in hex) for main memory address $(E201F)_{16}$?

- (A) F, 201 (B) E, 201 (C) E, E20 (D) 2, 01F

Block
Line



Cache memory organization part-1

$$\begin{aligned} \text{TAG} + 1 + 2 \\ (18 + 1 + 2) = 21 \text{ bits} \\ 256 \times 21 = 5376 \text{ bits} \end{aligned}$$

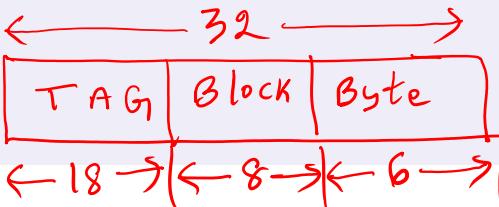
→ A 16 K Byte direct mapped write back cache is organized as multiple blocks, each of size 64 bytes. Processor generates 32 — bit addresses. Cache controller maintains the following information in addition to TAG for each cache block

1 valid bit, 2 modified bits

The total size of memory needed at the cache controller in bits is ?

- A. 5120
- B. 1024
- C. 1234
- D. ~~5376~~

$$\text{No. of Blocks} = \frac{16\text{KB}}{64\text{B}} = \frac{2^4}{2^6}$$



$$= \frac{8}{2} \text{ Blocks}$$

256

Cache memory organization part-1

$$\frac{4}{9} \times 0.9$$

$$\frac{205}{205} \times 2$$

$$1105$$

$$\begin{array}{r} 94.5 \\ 110.5 \\ \hline 205.0 \end{array}$$

→ A memory system has cache , main memory and secondary memory. Cache hit ration is 80 percent and main memory hit rate is 90 percent . If cache access time is 5 ns and main memory access takes 100 ns, Disk access takes 1000 ns, then what will be the effective access time ?

A. 40 ns

B. 400

C. 50

D. None of these

$$EAT = 0.8 \times 5 + 0.2 \left[0.9 [100 + 5] + 0.1 [1000 + 100 + 5] \right]$$

$$= 4 + 0.2 [0.9 \times 105 + 0.1 [1105]]$$

$$= 4 + 0.2 [94.5 + 110.5]$$

$$= 4 + 0.2 [205]$$

$$= 4 + 41 = \boxed{45}$$

Cache memory organization part-1

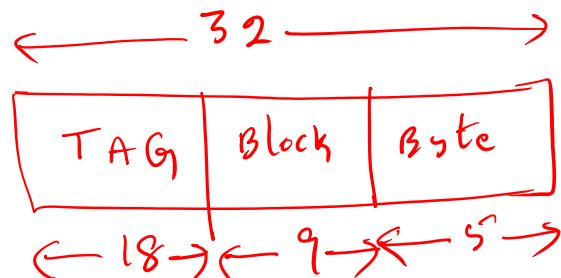
18

- Consider a machine with a byte addressable main memory of 2^{32} bytes divided into blocks of size 32 bytes. Assume that a direct mapped cache having 512 cache lines is used with this machine. The size of the tag field in bits is _____.

 2^{32} bytes

32 bytes

$$512 \times 32 = 2^9 \cdot 2^5$$

 2^{14} bytes16KBytes

Cache memory organization part-1

Cache Block No.

$$= \left(\frac{\text{Main. Mem.}}{\text{Block No.}} \right) \bmod \left(\frac{\text{No. of Blocks}}{\text{in Cache}} \right)$$

- Consider a direct mapped cache with 16 one word blocks, all initially empty, with respect to the given series of address reference

1, 4, 33, 5, 52, 4, 21, 5, 10, 19, 22, 10

Calculate the number of hits.

- A. 0
- B. 1
- C. 2
- D. 3

$$1 \bmod 16 = 1$$

$$4 \bmod 16 = 4$$

$$33 \bmod 16 = 1$$

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

Cache memory organization part-1

4 words (L₁)

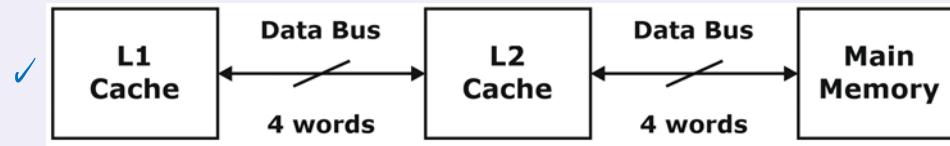
16 " (L₂)

20 n.s
2 m.s

22 n.s

H
2.

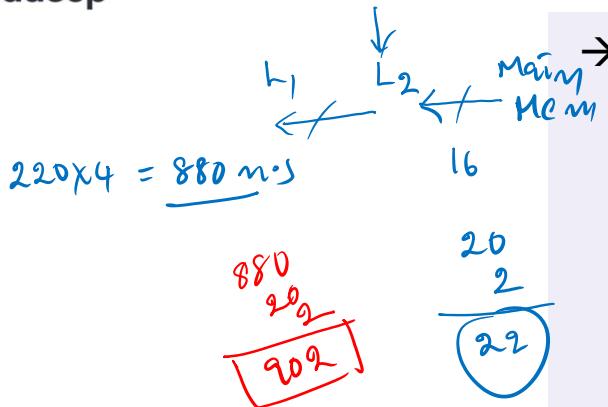
→ A computer system has an L1 cache, an L2 cache, and a main memory unit connected as shown below. The block size in L1 cache is 4 words. The block size in L2 cache is 16 words. The memory access times are 2 nanoseconds, 20 nanosecond and 200 nanoseconds for L1 cache, L2 cache and main memory unit respectively.



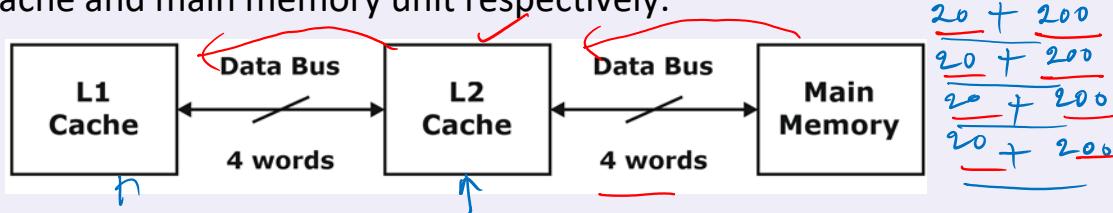
When there is a miss in L1 cache and a hit in L2 cache, a block is transferred from L2 cache to L1 cache. What is the time taken for this transfer ?

- A. 2 nanoseconds
- B. 20 nanoseconds
- C. 22 nanoseconds
- D. 88 nanoseconds

Cache memory organization part-1



A computer system has an L1 cache, an L2 cache, and a main memory unit connected as shown below. The block size in L1 cache is 4 words. The block size in L2 cache is 16 words. The memory access times are 2 nanoseconds, 20 nanosecond and 200 nanoseconds for L1 cache, L2 cache and main memory unit respectively.



When there is a miss in both L1 cache and L2 cache, first a block is transferred from main memory to L2 cache, and then a block is transferred from L2 cache to L1 cache. What is the total time taken for these transfers ?

- A. 222 nanoseconds
- B. 888 nanoseconds
- C. 902 nanoseconds
- D. 968 nanoseconds

$$\begin{array}{r}
 880 \\
 122 \\
 902
 \end{array}$$

Cache memory organization part-1

placement of Memory Blocks in Cache for fully Associative Mapping

4, 5, 9, 15, 12, 14, 1, 3, 17, 18, 20, 25, 8, 7, 9

LRU :- Least Recently used.

Replace the block that has not been used from the longest period of time.

0	4	17	2
1	5	18	16
2	9	20	15
3	15	25	
4	12	12	
5	14	7	
6	x	9	
7	3	1	

Cache memory organization part-1

H|W → An 8 KBytes direct mapped writeback Cache is organized as multiple blocks, the size of Cache block is 32 bytes. phy.Address has 30 bits. what are the TAG, Block-offset, Byte-offset bits respectively ?