



Computer Organization and Architecture

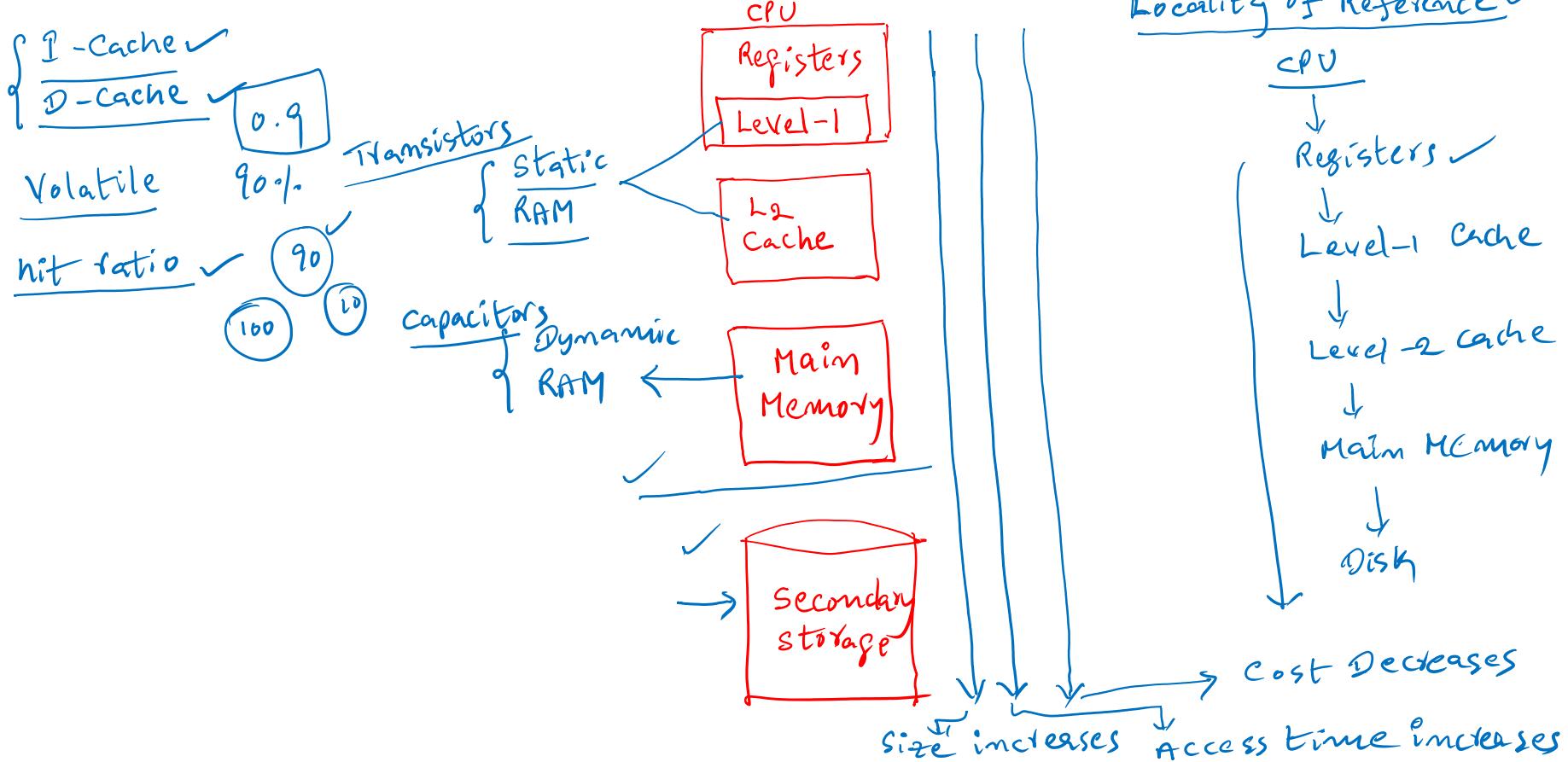
Memory Hierarchy

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



Memory Hierarchy



Memory Hierarchy

T_c [Cache Access]

T_m [Main Memory]

h (hit ratio in Cache)

Effective | Avg. Memory Access time

Two Level Memory system [Cache, Main Memory]

$$\text{EMT} = h * [T_c] + (1-h) [T_m + T_c]$$

$$h = 80\%$$

$$T_m = 200 \text{ n.s}$$

$$T_c = 25 \text{ n.s}$$

$$\begin{aligned} \text{EMT} &= 0.8 \times 25 + 0.2 [200 + 25] \\ &= 0.8 \times 25 + 0.2 \times 225 \end{aligned}$$

$$\begin{aligned} \text{Memory Access time} &= 20 + 45 \\ &= 65 \text{ n.s} \end{aligned}$$

→ The amount of time taken to finish a read request (or) write request

Memory Hierarchy

$\{$
 L₁-Cache
 L₂-Cache
 Main Memory

Three Level Memory Hierarchy

$$\begin{aligned}
 h_1 & [\text{Level-1}] \leftarrow T_{L1} [\text{Access time}] \rightarrow \text{Level-1} \\
 h_2 & [\text{Level-2}] \leftarrow T_{L2} [\text{Access time}] \rightarrow \text{Level-2} \\
 T_m & [\text{Access time of Main memory}]
 \end{aligned}$$

$$\text{Avg. Mean Access} = h_1(T_{L1}) + (1-h_1) \left[h_2(T_{L2} + T_{L1}) + (1-h_2) [T_m + T_{L2} + T_{C1}] \right]$$

Example

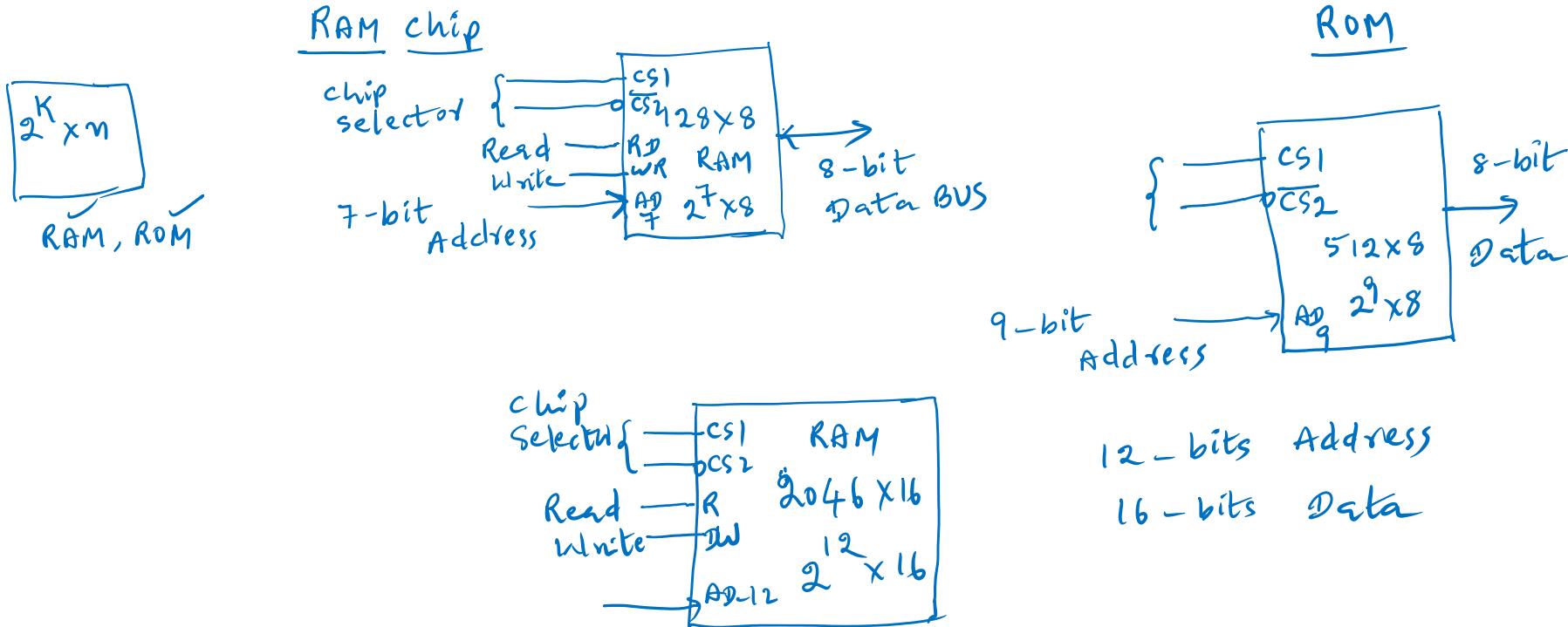
$$T_m = 100 \text{ n.s} \quad h_1 = 0.8 \quad \checkmark$$

$$\text{Level-1} = 10 \text{ n.s} \quad h_2 = 0.9 \quad \checkmark$$

$$\text{Level-2} = 20 \text{ n.s}$$

$$\begin{aligned}
 \text{E.M.T} &= 0.8 \times 10 + 0.2 [0.9 \times (20 + 10) + 0.1 \times 130] \\
 &= 8 + 0.2 [27 + 13] = 8 + 8 = 16 \text{ n.s}
 \end{aligned}$$

Memory Hierarchy



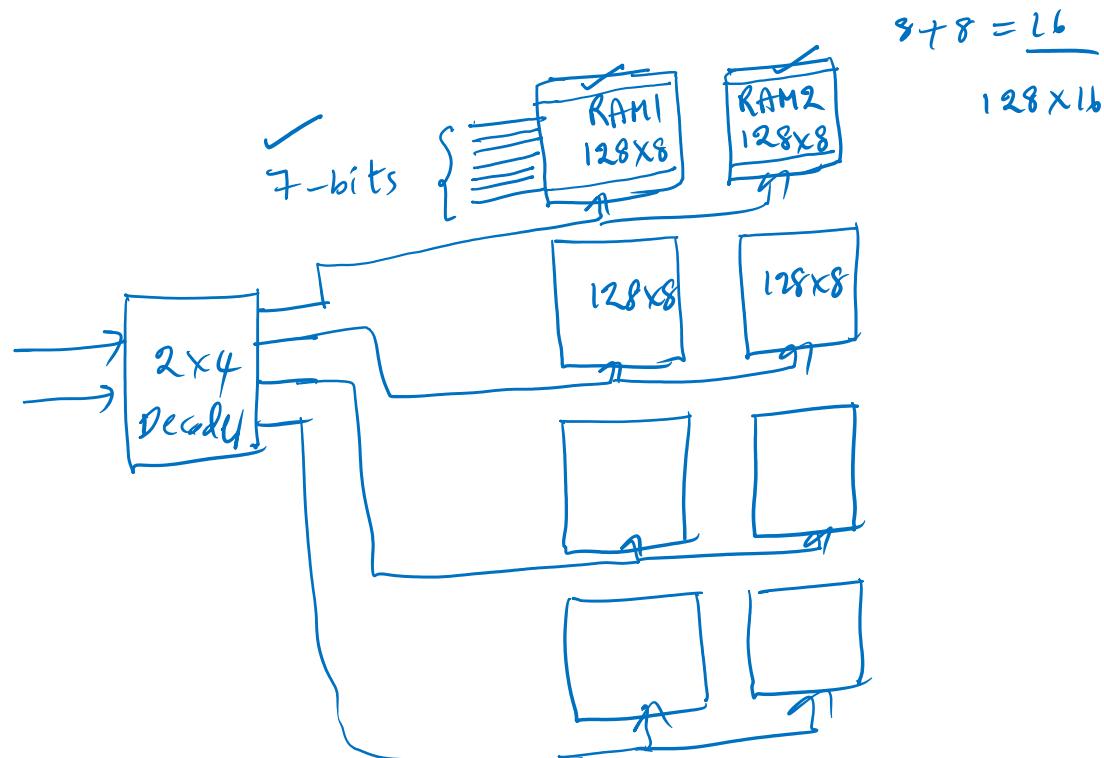
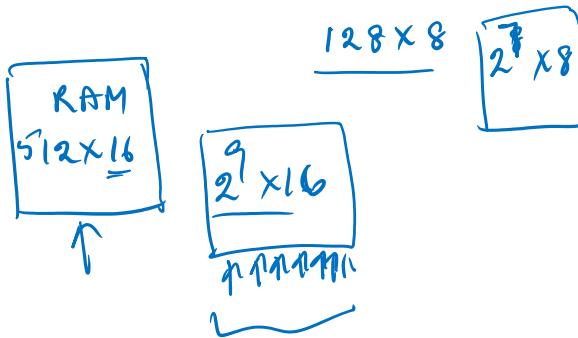
Memory Hierarchy

Function Table of RAM

<u>CS1</u>	<u>CS2</u>	RD	WR	Mem. function	status of Data Bus
1	0	X	X	Inhibit	High Impedance
RAM	ROM			"	"
1	0	0	0	"	"
1	0	1	X	Mem. Read	o/p's data from Memory
1	0	0	1	Mem. Write	i/p's data to memory
1	1	X	X	Inhibit	High Impedance

<u>CS1</u>	<u>CS2</u>	RD	WR	Mem. function	status of Data Bus
0	0	X	X	Inhibit	High Impedance
0	1	X	X	"	"
1	0	0	0	"	"
1	0	1	X	Mem. Read	o/p's data from Memory
1	0	0	1	Mem. Write	i/p's data to memory
1	1	X	X	Inhibit	High Impedance

Memory Hierarchy



Memory Hierarchy

Fetch Instructions

I-Cache, L₂-Cache, Main Memory

Fetch Operand

D-Cache, L₂-Cache, Main memory

- 1 The read access times and the hit ratios for different cache in a memory hierarchy are as given below.

Cache	Read access time (in nanoseconds)	Hit ratio
I-cache	2	0.8
D-cache	2	0.9
L2-cache	8	0.9

The read access time of main memory is 90 nanoseconds. Assume that the caches use the referred-word-first read policy and the write back policy. Assume that all the caches are direct mapped caches. Assume that the dirty bit is always 0 for all the blocks in the caches. In execution of a program. 60% of memory reads are for instruction fetch and 40% are for memory operand fetch. The average read access time in nanoseconds (up to 2 decimal places) is 4.72.

Memory Hierarchy

$$\begin{aligned}
 & \text{overall Avg} \\
 &= 0.6 \times 5.4 \\
 &\quad + 0.4 \times 3.7 \\
 &= 4.72
 \end{aligned}$$

Instruction Fetch

$$\begin{aligned}
 \text{Avg} &= 0.8(2) + 0.2[0.9[8+2] + 0.1[90+8+2]] \\
 &= 1.6 + 0.2[9+10] \\
 &= 1.6 + 3.8 = 5.4 \text{ n.s}
 \end{aligned}$$

Operand fetch

$$\begin{aligned}
 \text{Avg} &= 0.9(2) + 0.1[0.9[8+2] + 0.1[90+8+2]] \\
 &= 0.9 \times 2 + 0.1[9+10] \\
 &= 1.8 + 1.9 \\
 &= 3.7 \text{ n.s}
 \end{aligned}$$

$\frac{5.4 \times 6}{324}$ $\frac{3.7 \times 4}{148}$	$\frac{3.24}{1.48}$ $\frac{1}{4.72}$
--	---

Memory Hierarchy

Read = 50 n.s

miss

5 n.s for hit

2. Assume that for a certain processor , a read request takes 50 nanoseconds on a cache miss and 5 nanoseconds on cache hit. Suppose while running a program , it was observed that 80% of the processors read requests result in a cache hit. The average access time in nanoseconds is ?

- A. 25
- B. 20
- C. 14
- D. None of these

hit ratio = 0.8

$$\begin{aligned}\text{Avg} &= 0.8 \times 5 + 0.2 \times 50 \\ &= 4.0 + 10 \\ &= 14\end{aligned}$$

Memory Hierarchy

$$T_m = \frac{T_m}{5}$$

$$\text{EMT} = 50 \text{ n.s} \quad \left| \begin{array}{l} S_0 = \frac{2 T_m}{5} \\ T_m = \frac{50 \times 5}{2} \\ \boxed{50 + 0.2(80) = 60} \end{array} \right. \quad \underline{= 125 \text{ n.s}}$$

$$\left(\frac{T_m}{5} \right) = T_c \left(\frac{125}{1} \right) \text{ Q1}$$

- Q1. A two level memory system has a cache and main memory . Cache is having a hit ratio of 80%. A cache access time is 5 times faster than that of main memory access. If effective access time is 50 nanoseconds , what are the access times of cache and main memory respectively ?

- A. ~~125 and 25 nanoseconds~~
- B. 20 and 100 nanoseconds
- C. 25 and 125 nanoseconds
- D. None of these

$$\begin{aligned} S_0 &= 0.8 \times \frac{T_m}{5} + 0.2 \left[T_m + \frac{T_m}{5} \right] \\ &= 0.8 \frac{T_m}{5} + 0.2 \left[\frac{6 T_m}{5} \right] \\ &= 0.8 \frac{T_m}{5} + 1.02 \frac{T_m}{5} \end{aligned}$$

- Q2. If the effective access time is increased by 20 percent from 50 nanoseconds then what will be the change in the hit ratio ?

8/10

$$T_c = 25 \text{ n.s} \quad T_m = 125 \text{ n.s}$$

$\xrightarrow{x} 72$

$$\frac{60}{60} = x \times 25 + (1-x) [125 + 25]$$

$$60 = 25x + 150 - 150x$$

$$60 = 125x \quad , \quad x = \frac{60}{125} \times \frac{100}{5} = \frac{360}{125} = \frac{72}{25}$$

Memory Hierarchy

100 IF, 60 (OF)

40 (WO)

160 Mem. Reads

40 Writes

Total operations = 200.

$$\text{Avg. Mem. Access} = \frac{(224 + 112)}{200}$$

$$= \frac{336}{200} = 1.68$$

4. The memory access time is 1 nanosecond for a read operation with a hit in cache, 5 nanoseconds for a read operation with a miss in cache, 2 nanoseconds for a write operation with a hit in cache and 10 nanoseconds for a write operation with a miss in cache. Execution of a sequence of instructions involves 100 instructions fetch operations, 60 memory operand read operations and 40 memory operand write operations. The cache hit-ratio is 0.9. The average memory access time (in nanoseconds) in executing the sequence of instructions is ?

$$\cdot \frac{16 \times 1.4}{224} \quad \frac{328 \times 4}{212}$$

Reads

$$\text{Avg} = 0.9 \times 1 + 0.1 \times 5 = 0.9 + 0.5 = 1.4 \text{ n.s}$$

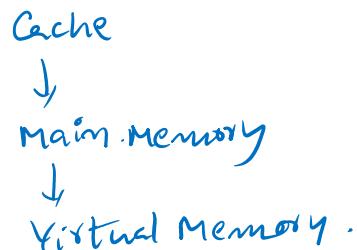
For 160 reads, $160 \times 1.4 = 224 \text{ n.s}$

writes

$$\text{Avg} = 0.9 \times 2 + 0.1 \times [10] = 1.8 + 1 = 2.8$$

For 40 writes, $40 \times 2.8 = 112 \text{ n.s}$

Memory Hierarchy



→ A 3-level memory system has a Cache, a main memory and virtual memory. Cache access takes 2 cycles, Main memory access takes 150 cycles, and virtual memory takes 1 Latch cycles. If Cache hit rate is 80% and main memory hit rate is 90%, what is the average memory access time? [In cycles]

Sol

Avg. Mem. Access time

$$= 0.8 \times 2 + 0.2 [0.9 \times [150 + 2] + 0.1 \times [100000 + 150 + 2]]$$

$$= 0.8 \times 2 + 0.2 [0.9 \times 152 + 0.1 [100152]]$$

$$= 1.6 + 0.2 [136 + 10015.2]$$

$$= 2031.8$$