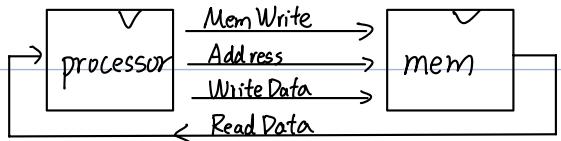
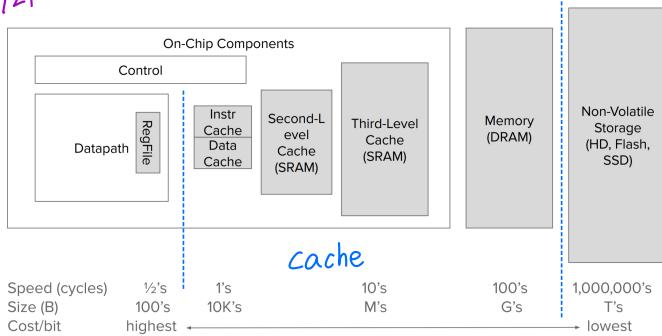


Mem physically separate from processor. Slow :/



To keep up: hierarchy (speed vs size)

11/21



Temporal locality if data recently used → likely used in future

Spacial ~ nearby data used in future

Performance Hit - data found in that level of mem.

Miss - X → next level

$$AMAT_i = \text{Access Time}_i + \text{Miss Rate}_i \cdot AMAT_{i+1} \quad (\text{recursive hierarchy})$$

E. 2000 S/L, 1250 in caches 750 not,  $\text{Acc Time}_{\text{cache}} = 1 \text{ cycle}$ ;  $\text{Acc Time}_{\text{mem}} = 100 \text{ cycle}$

$$1 + \frac{750}{2000} \cdot 100 = 38.5 \text{ cycles}$$

or neighboring

Cache: recent data (too old → replaced)

addressed w/ simple address hash

Def. Capacity (CC): # bytes = Bb

$$\# \text{ blocks } B = SN$$

$$\text{Block size } b \text{ (byte/block)}$$

$$\# \text{ sets } S$$

$$\text{Deg. of associativity } N$$

addr  $\xrightarrow{\text{hash}}$  one set, index set, check all ways

4:2

Hash: pull a few bits from addr.

8 words (B) (meta not included)

8

1 word/block

8

1

direct map

col: ways (1)

block

row:

sets (8)

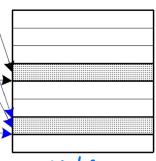
(if 1 set, Bways,  
fully associative)

Address

11...11111100	mem[0xFF..FC]
11...11111000	mem[0xFF..F8]
11...11110100	mem[0xFF..F4]
11...11110000	mem[0xFF..F0]
11...11101100	mem[0xFF..EC]
11...11101000	mem[0xFF..E8]
11...11100100	mem[0xFF..E4]
11...11100000	mem[0xFF..E0]
⋮	⋮
00...00 001000	mem[0x00..24]
00...00100000	mem[0x00..20]
00...00011100	mem[0x00..1C]
00...00011000	mem[0x00..18]
00...00010100	mem[0x00..14]
00...00010000	mem[0x00..10]
00...00001100	mem[0x00..0C]
00...00001000	mem[0x00..08]
00...00000100	mem[0x00..04]
00...00000000	mem[0x00..00]

(most significant mem)

Set Number
7 (111)
6 (110)
5 (101)
4 100
3 (011)
2 (010)
1 001
0 (000)



cache

(mod 4)

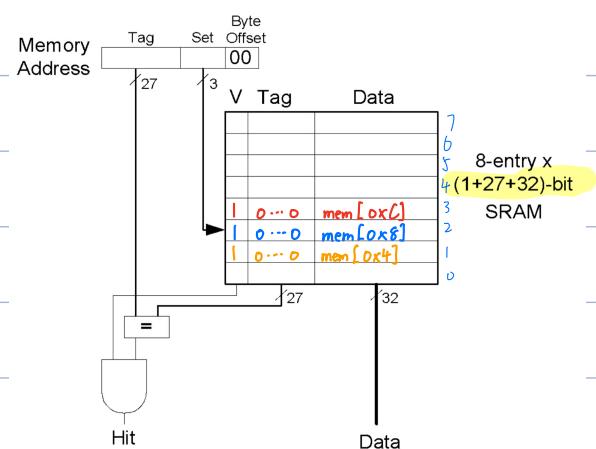
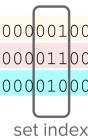
V: valid bit,  $\checkmark$ ; Tag: remainder of addr.

Hit = V and (Tags =)

E. load 0x4, 0xC, 0x8 5 times

```

addi $t0, $0, 5
loop: beq $t0, $0, done
lw $t1, 0x4($0) 00000000 00000000 00000000 00000100
lw $t2, 0xC($0) 00000000 00000000 00000000 000001100
lw $t3, 0x8($0) 00000000 00000000 00000000 000001000
addi $t0, $t0, -1
j loop
done:
    
```



First time miss. Once in cache → saved → hit

3.1 Xs

3. (5-1)  $\checkmark$ s  $\rightarrow$  20% miss rate

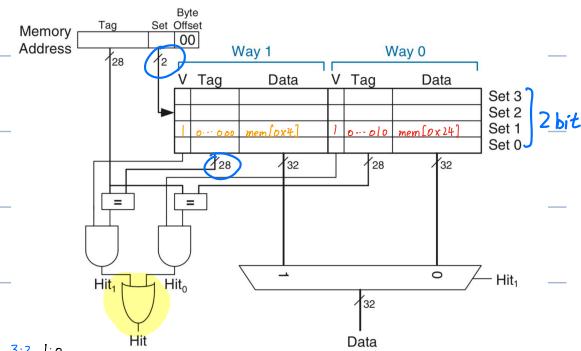
If addr. are 0x4, 0x24 → conflict misses  $\rightarrow$  100% miss : (

2-way associative cache ( $8 \times 1 \rightarrow 4 \times 2$ ) idx: 3  $\rightarrow$  2

E. load 0x4, 0x24 5 times

Fully-asso.  $\rightarrow$  min. conflict miss, slow, \$\$

more read



Spatial block size  $\uparrow$

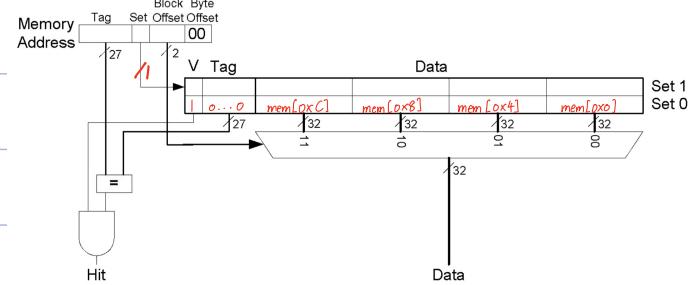
E.  $b = 4$  words  $\rightarrow B = \frac{C}{b} = 2$  blocks

$\log_2 4$

$3b \rightarrow 1$  set + 2 block offset

E. 0x4, 0x8, 0xC

only 0x4 L miss  $\rightarrow$  loads 0x0 ~ 0xC  $\rightarrow$  all hit!  $\rightarrow$  miss rate = 1/15



Miss 1. Compulsory: 1<sup>st</sup> time access  $\rightarrow 0$

2. Capacity: too much data to cache  $\leftarrow C$

3. Conflict: map to same location  $\leftarrow C, N$

but as CT, AMAT ↴

