



TED ÜNİVERSİTESİ

Computer Organization – CMPE361

Department of Computer Engineering
TED University- Fall 2023

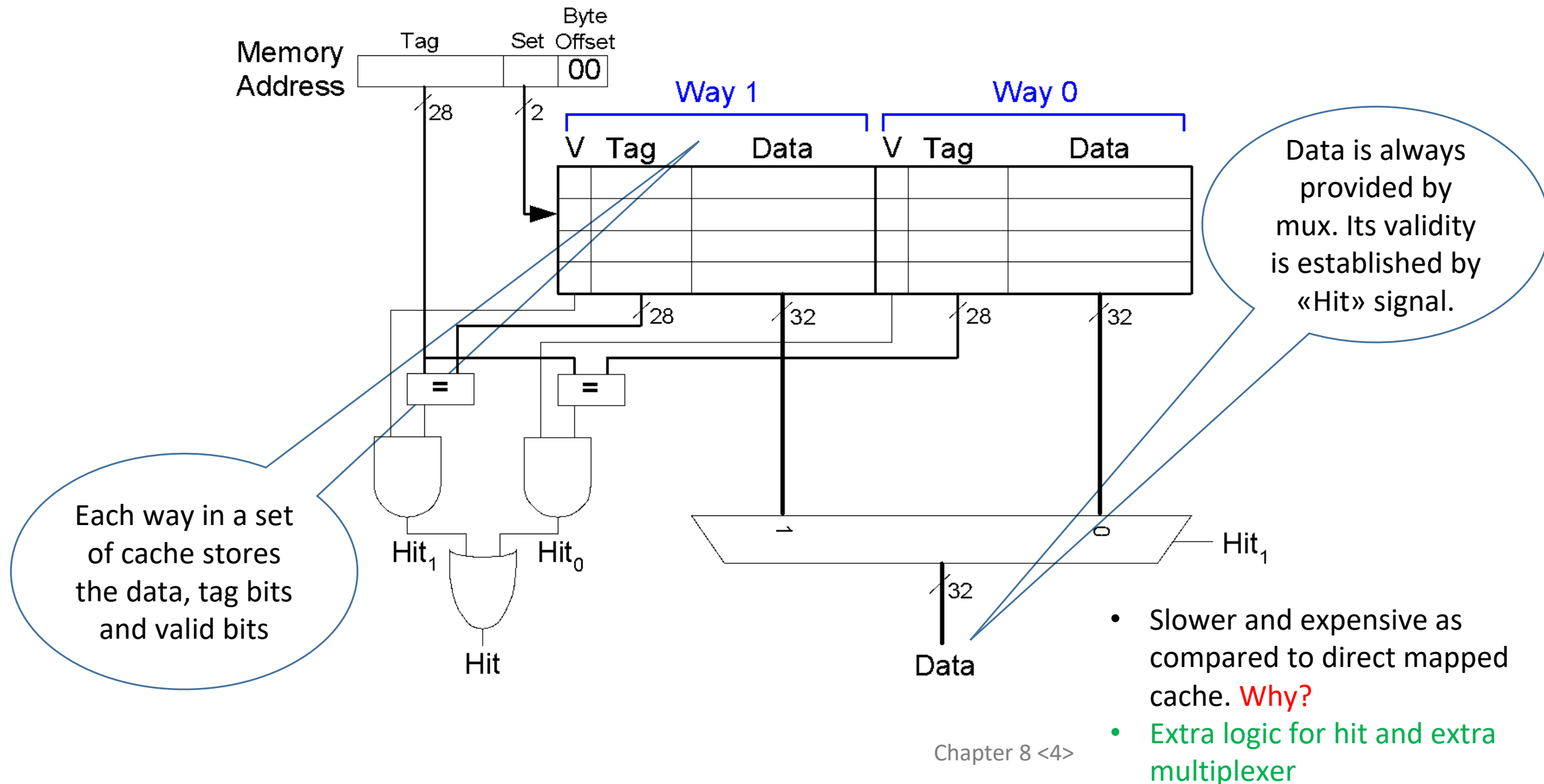
Memory Systems 3- Caches

These Slides are mainly based on slides of the text book (downloadable from the book's website).

How to reduce conflict misses of Direct Mapped Cache?

- Increase the number of blocks in a set.
- This is known as increasing the associativity of cache.
- N-way associative cache has N blocks in each set.
- Each of memory addresses still maps to only one set (like a direct mapped cache).
 - However, it can map to any of the blocks inside that set.

N-Way Set Associative Cache



- Slower and expensive as compared to direct mapped cache. **Why?**
- **Extra logic for hit and extra multiplexer**

N-Way Set Associative Example

MIPS assembly code

```

        addi $t0, $0, 5
loop:   beq  $t0, $0, done
        lw   $t1, 0x4($0)
        lw   $t2, 0x24($0)
        addi $t0, $t0, -1
        j    loop
done:

```

Miss Rate = ?

Way 1			Way 0			
V	Tag	Data	V	Tag	Data	
0			0			Set 3
0			0			Set 2
0			0			Set 1
0			0			Set 0

N-Way Set Associative

MIPS assembly code

```

        addi $t0, $0, 5
loop:   beq  $t0, $0, done
        lw   $t1, 0x4($0)
        lw   $t2, 0x24($0)
        addi $t0, $t0, -1
        j    loop
done:

```

N=2

Miss Rate = 2/10
= 20%

Associativity reduces
conflict misses

Problem:

If both ways of a set are full, and a new block is to be placed in the same set, which one of two (already existing blocks in the set) shall be evicted? Is a valid problem to be addressed later!

Way 1			Way 0			
V	Tag	Data	V	Tag	Data	
0			0			Set 3
0			0			Set 2
1	00...10	mem[0x00...24]	1	00...00	mem[0x00...04]	Set 1
0			0			Set 0

Fully Associative Cache

V	Tag	Data	V	Tag	Data	V	Tag	Data	V	Tag	Data	V	Tag	Data	V	Tag	Data	V	Tag	Data

One set containing all blocks

Fewest conflict misses

Expensive to build

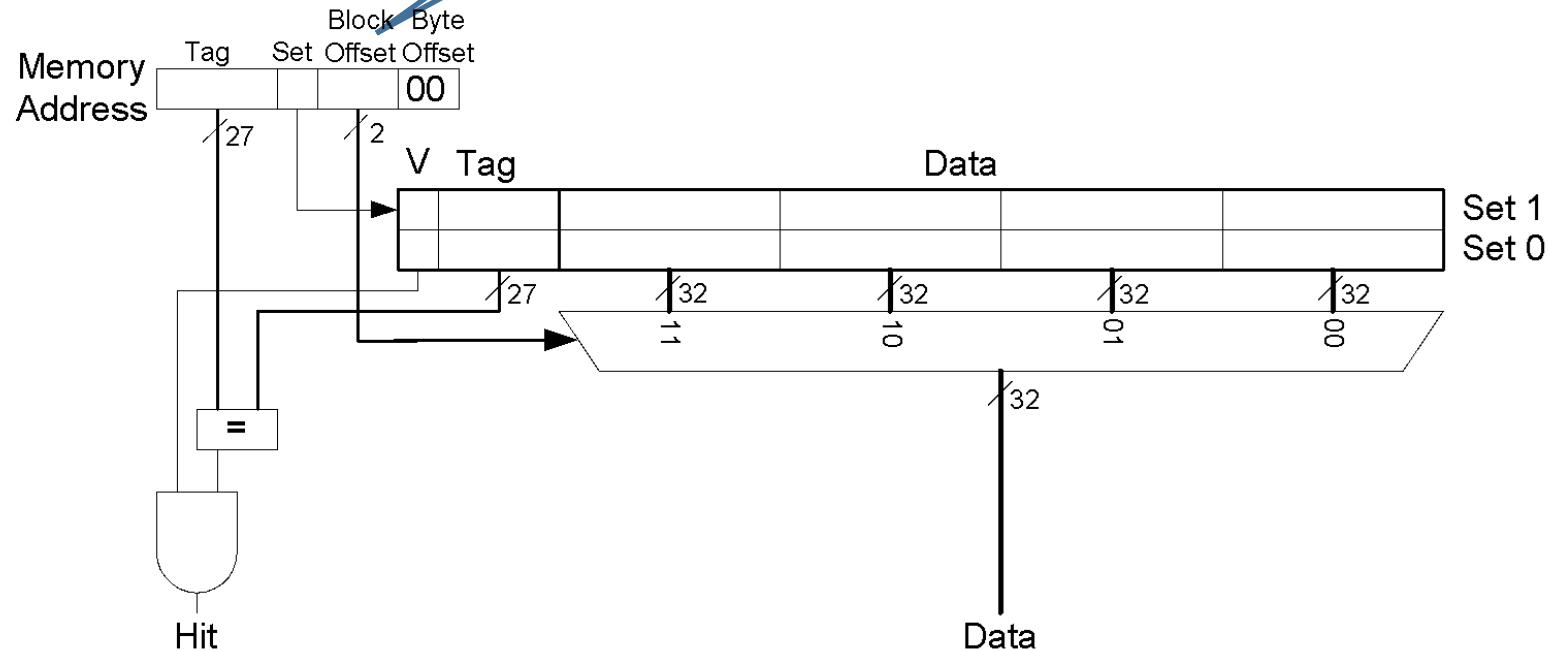
How to take advantage of spatial locality in caches?

- When reading a Word from memory into cache, the whole block to which the required Word belongs is brought **using block boundaries**.
- Suppose Word address is 0x00000028, then a block (of 4 words) will be read starting from address 0x00000020. (b'...1 0 00 00)
- Another example, Word address is 0x35F4024C, what is the starting address of the corresponding block? (b'...10 0 11 00)= 0x35F40240

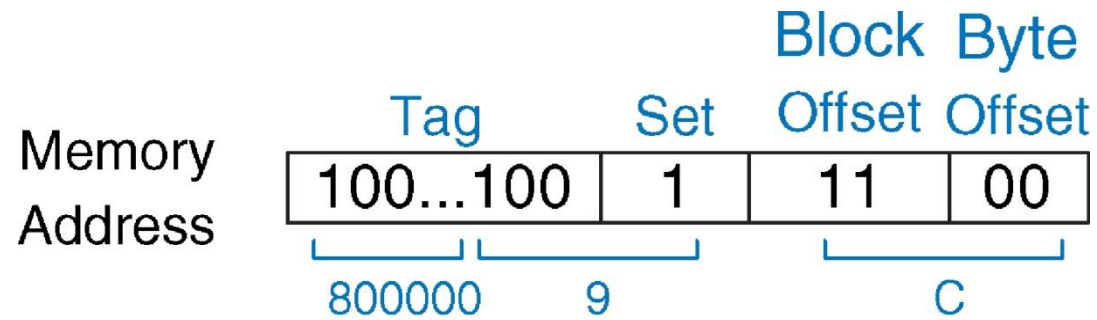
- Increase block size:

- Block size, **$b = 4$ words**
- $C = 8$ words
- Direct mapped (1 block per set)
- Number of Sets, **$N = 2$** ($C/b = 8/4 = 2$)

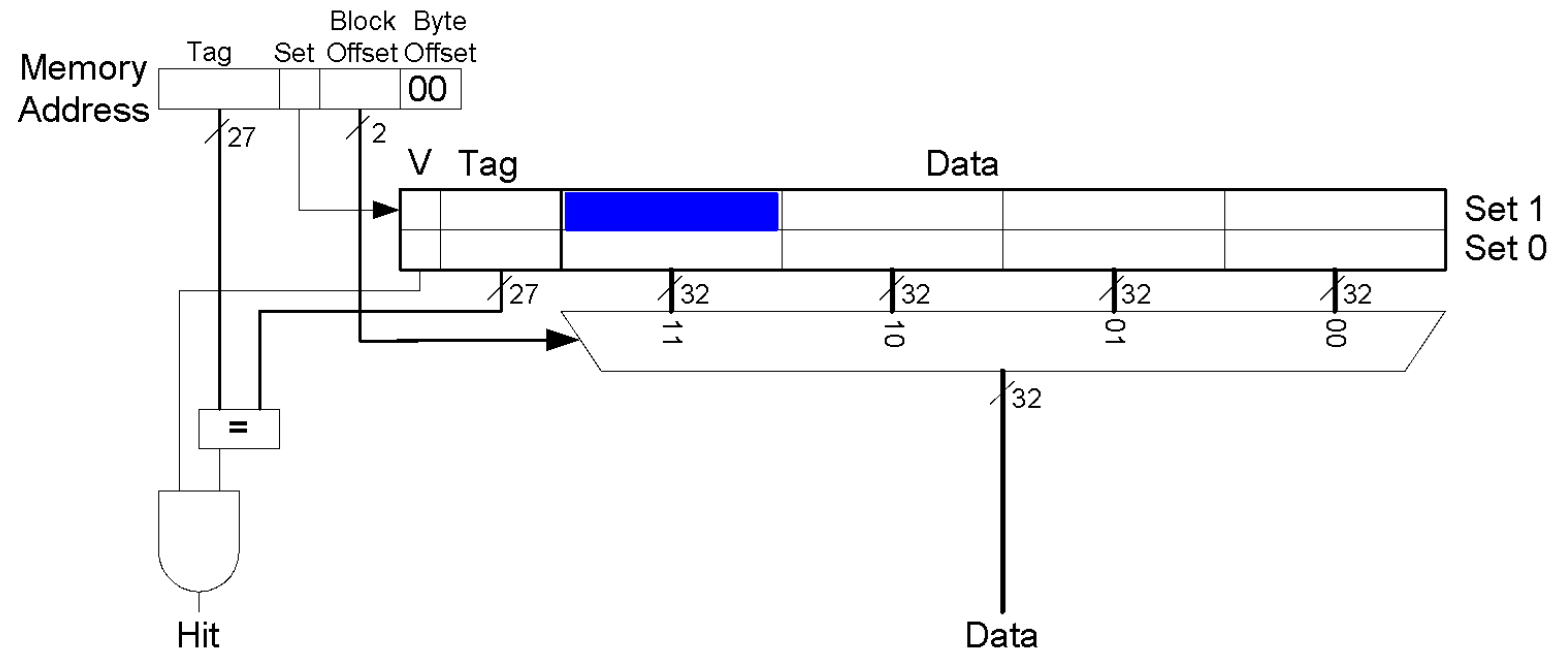
A new field is introduced to select a word inside a block.
Number of bits in BlockOffset field= $\log_2 4 = 2$ bits



Example: Cache with Larger Block



© 2007 Elsevier, Inc. All rights reserved



Example: Direct Mapped Cache with large block

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

Miss Rate = ?

Direct Mapped Cache Performance with larger block size

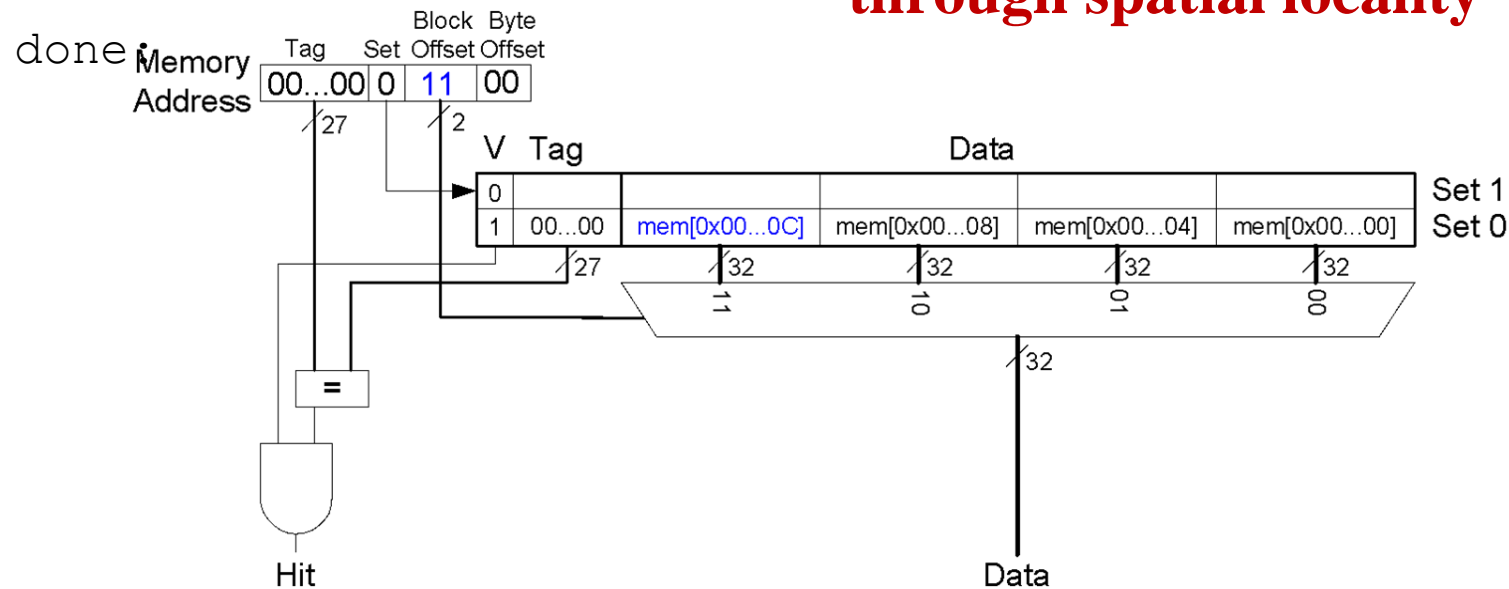
```

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

```

Miss Rate = 1/15
= 6.67%

Larger blocks
reduce compulsory misses
through spatial locality



Cache Organization Summary

- Capacity: C
- Block size: b
- Number of blocks in cache: $B = C/b$
- Number of blocks in a set: N
- Number of sets: $S = B/N$

Organization	Number of Ways (N)	Number of Sets ($S = B/N$)
Direct Mapped	1	B
N-Way Set Associative	$1 < N < B$	B / N
Fully Associative	B	1