

CS 223: Computer Architecture & Organization

Cache Memory



J. K. Deka

Professor

**Department of Computer Science & Engineering
Indian Institute of Technology Guwahati, Assam.**

Cache Mapping

- Direct Mapping
- Associative Mapping
- Set Associative Mapping
 - K way Set Associative Mapping

Mapping: Example

- A block set associative cache consists of a total of 64 lines divided into 4-line sets. The main memory contains 4096 blocks, each consisting of 128 words.
 - What is the size of main memory and cache memory

Mapping: Example

- A block set associative cache consists of a total of 64 lines divided into 4-line sets. The main memory contains 4096 blocks, each consisting of 128 words.
 - How many bits are there in a main memory address
 - How many bits are there in each of the TAG, SET and WORD fields.

Replacement Algorithms

- Direct mapping
 - No choice
 - Each block only maps to one line
 - Replace that line

Replacement Algorithms

- Associative & Set Associative
 - Hardware implemented algorithm (speed)
 - Least Recently used (LRU)
 - e.g. in 4 way set associative
 - Which of the 4 blocks is lru?
 - First in first out (FIFO)
 - replace block that has been in cache longest
 - Least frequently used
 - replace block which has had fewest hits
 - Random

Replacement Algorithms

- Least Recently Used (LRU)
 - Program usually stays in localized area for a reasonable period of time.
 - There is a high probability that the blocks that have been referenced recently will be referenced again soon.
 - When a block is to be overwritten, it is sensible to overwrite the one that has gone the longest time without being referenced.
 - This block is called the **Least Recently Used (LRU) block** and the technique is called the LRU replacement policy.

Least Recently Used (LRU)

- Consider four-line set in a set-associative cache
- Control bits:
 - TAG bits
 - 2-bit counter for each line(to track the LRU block)
 - d_bit: dirty bit
 - f_bit: occupied bit
- Initially reset all the counters, d_bit and f_bit

Least Recently Used (LRU)

- A cache hit occurs:
 - set the counter value to 0 of this cache line
 - for other counters, if the value is less than the referenced line, increment the counter value provided f_bit is 1.
 - otherwise, do not change the counter value

Least Recently Used (LRU)

- A cache miss occurs:
 - Set is not full
 - Set is full
- Set is not full:
 - set the counter value to 0 of the cache line
 - set f_bit to 1 and bring the data to this cache line
 - increment the counter value of other lines whose f_bit is 1

Least Recently Used (LRU)

- Set is full:
 - the line with highest counter value is removed; write back if d_bit is 1
 - new block is transferred to this line
 - reset d_bit and the counter
 - other counter values are incremented by 1

Example

- Consider the following code segment

```
for (i=0; i<100; i++){  
    for (j=0; j<100; j++)  
        B[i,j] = A[i,j];}
```

- Cache organization:
 - 16 lines/blocks in the cache
 - block size is 512
- Data in Main memory:
 - block0 contains i, j, etc
 - block1 onward: Array A
 - block25 onward: Array B

Example

Array stored in memory:

- Row Major Order
- Column Major Order

Direct Mapping

- Mapping function
 - $\text{cache line} = i \bmod 16$ (for i th block)
 - line 0 : blocks 0, 16, 32,
 - line 1 : blocks 1, 17, 33,
- Mapping for Array A
- Mapping for Array B
- Find:
 - Cache hits
 - Cache misses

2-way set associative mapping

- Mapping function
 - $i = j \bmod v$ (v is number of set = 8)
 - $m = v * k$ ($k = 2$, lines in a set)
 - set0: line 0, 1: blocks 0, 8, 16, 24, 32, ...
 - set1: line 2, 3: blocks 1, 9, 17, 25, 33,
- Mapping for array A
- Mapping for array B
- Replacement policy:
 - FIFO (cache hits and cache misses)
 - LRU (cache hits and cache misses)

2-way set associative mapping

- Mapping for array A
- Mapping for array B

Reference

Computer Organization and Architecture –
Designing for Performance
William Stallings, Seventh Edition

Chapter 04: Cache Memory

Computer Organization
Hamacher, Vranesic and Zaky, Fifth Edition

Chapter05: Page No.: 314 - 329