CENG3420 Homework 4

Due: 23:59 PM, Apr. 24, 2022

All solutions should be submitted to the blackboard in the format of **PDF/MS Word**.

Q1 (10%) There are many notations used to describe a cache. The following two tables summarize the symbols we refer to describe a cache:

Parameter	Description
$S = 2^s$	Number of sets
E	Number of blocks
$B = 2^b$	Block size (bytes)
$m = \log_2(M)$	Number of main memory address bits

Table 1: Fundamental parameters.

Parameter	Description
$M = 2^m$	Maximum number of unique memory addresses
$s = \log_2(S)$	Number of set index bits
$b = \log_2(B)$	Number of block offset bits
$t = m - (s + b)$	Number of tag bits
$C = B \times E \times S$	Cache size not including overhead (valid and tag bits)

Table 2: Derived quantities.

Complete the form for different caches below:

Cache id	m	C	B	E	S	t	s	b
1.	32	1024	4	1				
2.	32	1024	8	4				
3.	32	1024	32	32				

Q2 (20%) A processor has a 32-bit memory address space (i.e. 32-bit addresses). The memory is broken into blocks of 32 bytes each. The computer also has a cache capable of storing 16K bytes.

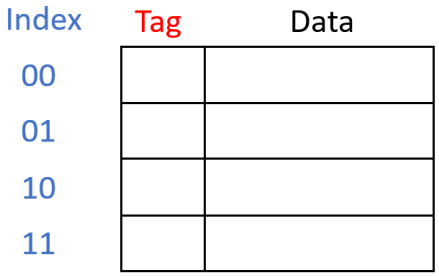
1. How many blocks can the cache store?
2. Assuming the cache uses **direct-mapping**, how many bits are there in each of the Tag, Block, and Byte offset fields fields of the address.

Tag	Block	Byte offset

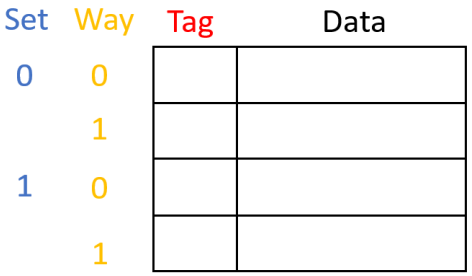
3. Assuming the cache uses **2-way set associative mapping**, how many bits are there in each of the Tag, Set, and Byte offset fields fields of the address.

Tag	Set	Byte offset

4. **Direct mapping** v.s. **2-way set associate mapping**.
   Consider the following two empty caches, calculate cache hit rates for the reference word addresses: (a) "0, 4, 0, 4, 0, 4, 0, 4"; (b) "0, 3, 0, 3, 0, 3, 0, 3".



Direct Mapping



2-Way Set Associative

Q3 (20%) In the following questions, start from an empty cache and give the contents of the cache after the following sequence of memory references (addresses are hexadecimal numbers): A0, F1, FF, 35, C8, 89, FE, 88, A1, A2, A3, A9, 99, 80, 83.

1. A block=2, 4-way cache with the LRU replacement.

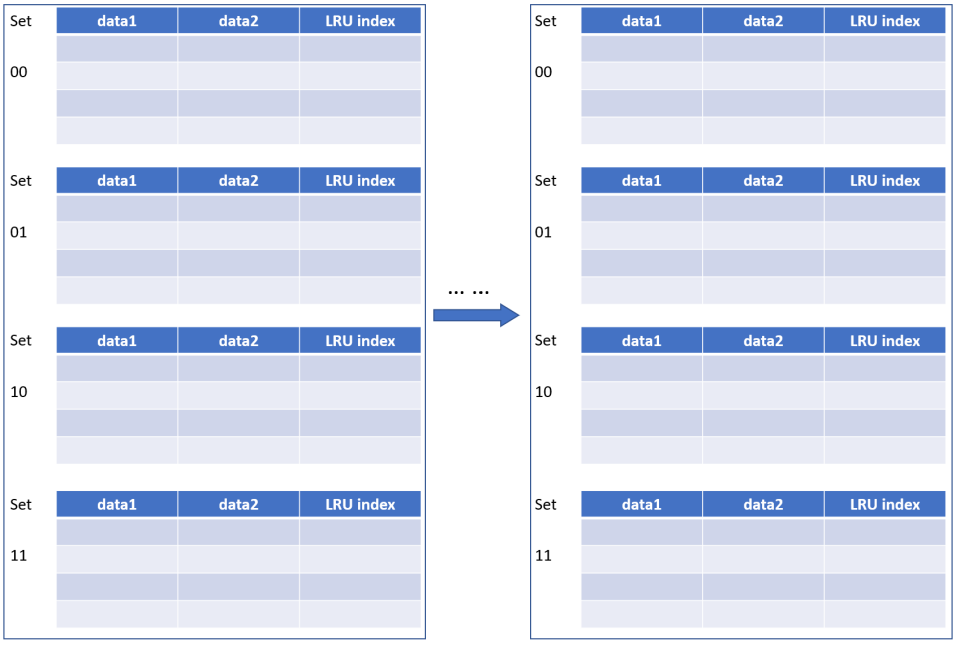


Figure 1: Examples of the contents of the block=2, 4-way cache with the LRU replacement.

2. A block=4, 2-way cache with the FIFO replacement.

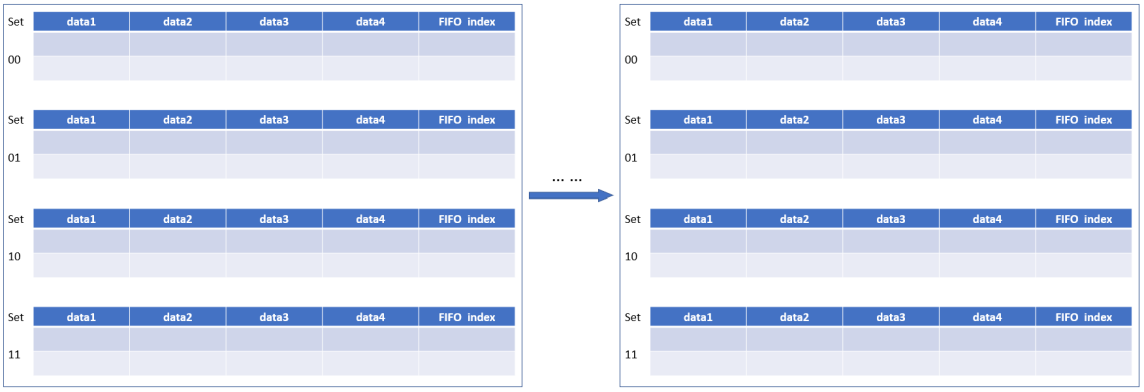
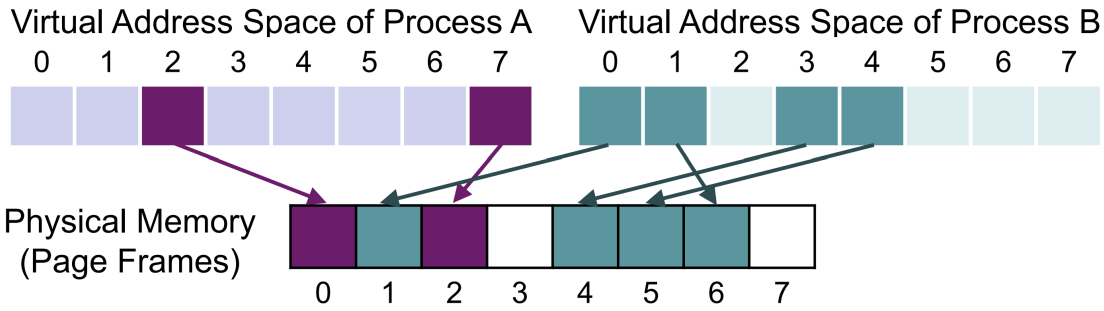


Figure 2: Examples of the contents of the block=4, 2-way cache with the FIFO replacement.

Q4 (10%) Suppose the access times to the cache and the main memory are 50 ns and 200 ns respectively. When the CPU executes a program, it accesses the cache 2000 times and main memory 50 times. Calculate the hit rate and access efficiency of this cache-main memory system.

Q5 (10%) Please draw the page tables for processes A and B:



Q6 (20%) Problems in this exercise refer to the following sequence of instructions, and assume that it is executed on a 2-issue RISC-V:

```
LOOP: lw   x31, 0(x20)     # x31 = some value
      add  x31, x31, x21   # add scalar in x 21
      sw   x31, 0(x20)     # store result
      addi x20, x20, -8    # decrement pointer
      blt  x22, x20, LOOP # branch if x22 < x20
```

1. Using the scheduled instruction to calculate IPC (instructions per clock cycle).
2. Suppose we have four registers (x28, x29, x30, x31), please design a solution to unroll the loop for better IPC.

Q7 (10%) Assume we have a program where 10% of the execution time is purely sequential and that the rest of the execution time can be improved by parallelization. For the part of the code that can be parallelized, each core gives only 80% improvement. For instance, 5 cores give $5 \times 80\% = 4$ times improvement.

1. Create a speedup chart, showing speedup on the Y-axis and the number of cores on the X-axis. Show the graph for 25 to 200 cores, for instance by plotting with 25 cores interval.
2. What is the maximal speedup that can be achieved regardless how many cores we add?