Istanbul Technical University
Department of Computer Engineering

BLG 322E – Computer Architecture
Assignment 5

Due Date: 03.06.2020, **Wednesday,** 23.59.

QUESTION 1 (50 points):

In a computer system, the address bus and data bus are **20 bit**s and **8 bit**s wide, respectively. The system includes a cache memory that can hold **4 KB** data.

Suppose you are running a program with the following data access pattern. The hexadecimal numbers represent the addresses of data accessed consecutively. The pattern is executed only once. In the beginning, the cache memory is empty.

$00050, $0005C, $01052, $00057, $01054

a) Data transfer between main memory and cache is in blocks of **16 bytes**. The cache control unit uses **direct mapping**.

i) Into what fields does the cache control unit divide the physical address? How many bits are there in each field? Explain.
ii) For each address in the given data access pattern, write which event (a **hit** or a **miss**) occurs in cache memory.

b) Data transfer between main memory and cache is in blocks of **8 bytes**. The cache control unit uses **2-way set associative mapping**.

i) For each address in the given data access pattern, write which event (a **hit** or a **miss**) occurs in cache memory.

QUESTION 2 (50 points):

A single-CPU computer system has

- **512K words** of main memory and
- cache memory that can hold **16 K words** of data.

Data transfer between main memory and cache is in blocks of **16 words**. The cache control unit uses *set associative mapping*. The cache memory contains **256 sets**.

When necessary, **FIFO** is used as the replacement algorithm. Assume that Frame **i** is older than Frame **i+1** in each set. Cache memory is used only for data, not for instructions.

The CPU runs the piece of pseudocode given on the right. Twenty elements of array A are copied into array B, and ten elements of array A are copied into array C. Each element is one word. The starting addresses of the arrays are given below:

A: $0008E
B: $00119
C: $00212

Assume that cache memory is full, but arrays A, B, and C are not in cache at the beginning of the program.

```
LOOP1:
For i = 0 to 19
   B[i] = A[i];
End of For
LOOP2:
For i = 0 to 9
   C[i] = A[2*i];
End of For
```

The following questions (a and b) can be answered independently.

a) For write operations, **Write Through** (WT) with **NO Write Allocate** (NWA) is used.

i) Which set/frames of cache memory does the cache control unit place arrays A, B, and C into? Give your answers in decimal. (Example: "**set number: 73, frame: 0**" or "**set number 85, frame: 1**")
ii) How many read misses, read hits, write misses, write hits, write-to-main-memory operations, and block transfers occur during the run of the given loop?
iii) Explain the reason for each block transfer during the run of the given program.

b) For write operations, **Simple Write Back** (SWB) with **Write Allocate** (WA) is used.

i) How many read misses, read hits, write misses, write hits, write-to-main-memory operations, and block transfers occur during the run of the given loop?
ii) Explain the reason for each block transfer during the run of the given program.

Note: If you have questions about the homework, you may contact the research assistants of the course (ozcelikfu@itu.edu.tr).