Student ID:
First Name and Last Name:
Signature:

İTÜ Computer Engineering Department
08.06.2017
Assoc.Prof. Feza BUZLUCA
Asst. Prof. Sanem KABADAYI

COMPUTER ARCHITECTURE FINAL EXAMINATION

Regulations:

1. The exam duration is 110 minutes.
2. You may not ask the proctors any questions during the exam.
3. Cell phones are prohibited on the desk; they must be **switched off**.
4. Any cheating or any attempt to cheat will be subject to the University disciplinary proceedings.

QUESTION 1: (25 Points)

We design a pipeline P to perform task **T** on elements of an array.
The speedup achieved by pipeline P if the array has 10 elements (n=10) is $S_{n=10} = 2$.
The speedup achieved by pipeline P if the array has an infinite number of elements is $S_{n\to\infty} = 3$.

a) What is the theoretical maximum speedup $S_{max(theoretical)}$ for this pipeline? (10 p)

$S_{n=10} = 10xtn/((k+9)xtp) = 2$, $S_{n\to\infty} = tn/tp = 3 \rightarrow k=6$

$S_{max(theoretical)} = 6$

b) If the duration for executing the task on only the first element using the pipeline P is $\mathbf{T}_1 = 180$ ns, what is the cycle time tp of the pipeline P? (10 p)

$T_1 = kxtp = 180$ ns

tp= 30ns

c) We redesign pipeline P to have a higher number of stages (layers) but the cycle time tp remains the same. How does $S_{n\to\infty}$ (speedup achieved by the pipeline if the array has an infinite number of elements) change? Choices: $S_{n\to\infty}$ decreases, increases, or remains the same. (5 p)

$S_{n\to\infty} = tn/tp$ Since tn and tp do not change, $S_{n\to\infty}$ remains the **same**.

Student ID:
First Name and Last Name:
Signature:

QUESTION 2: (20 Points)

In a RAID system, the total number of disks is **4** (all disks in the system).

The following questions can be answered independently.

a) How many disks can be accessed simultaneously (at the same time) for the read operation in the RAID systems given below? Write the number in the given blank space for each system below. (10p)

RAID 0: 4 disks

RAID 1: 2 disks

RAID 4: 3 disks

RAID 5: 4 disks

RAID 6: 4 disks

Note: For the read operations, the parities are not necessary.
Parity disks are accessed for two reasons:
1. In write operations, to construct the parity information
2. In case of a disk failure, to recover data
Since the disks in RAID 5 and RAID 6 are not synchronized and parities are distributed, 4 disks can be accessed at the same time.

b) In what ways is RAID 5 superior to (better than) RAID 4? Write "YES" or "NO" next to each of the five metrics below. (10p)

Cost: NO

Average read access time: NO

Worst-case read access time: NO

Average write access time: YES

Worst-case write access time: NO

Since the parity strips are distributed across all disks in RAID 5, the possibility for a write penalty is lower than in RAID 4. Therefore, the average write access time is shorter in RAID 5.
However, in the worst case where data are written to the same physical disk or to different disks with the common parity disk, both systems have write penalties.

Student ID:
First Name and Last Name:
Signature:

QUESTION 3: (30 Points)

A computer system has **64 KxWords** of main memory and a cache memory that can hold **4 KxWords** of data. Data transfers between main and cache memories are performed using blocks of **16 words**. The cache control unit uses the *two-way set associative mapping* technique where each set contains **two frames**.
In necessary cases, **FIFO** is used as the replacement algorithm.
Cache memory is used only for data, not for instructions.
The following questions (a and b) can be answered independently.

a) For write operations, the **Write Through** (WT) with **NO Write Allocate** (NWA) method is used.
Assume that the cache memory is empty at the beginning. The CPU runs the piece of pseudo code given on the right. Ten elements of array A are copied to array B. Each element is one word. The starting addresses of the arrays are given below:

```
For i = 0 to 9
    B[i] = A[i];
End of For
```

A: $083C
B: $0030

i) Which set/frames of cache memory does the cache control unit place arrays A and B into? Write your answers as decimal numbers. (Example: "**set number: 73, first frame**" or "**set number 85, second frame**.") (10 p)

A: set number: 3, first frame and set number: 4, first frame
Note: Because of the NO Write Allocate (NWA) method, array B is not loaded to the cache.

ii) How many read misses, read hits, write misses, write hits, and block transfers occur during the run of the given loop? Write your answers in the boxes below: (10 p)

Number of Read misses:	2	Number of Read hits:	8
Number of Write misses:	10	Number of Write hits:	0

Number of Block transfers: 2

b) For write operations, the **Flagged Write Back** (FWB) with **Write Allocate** (WA) method is used.
Assume that the cache memory is empty at the beginning. The CPU runs the piece of pseudo code given on the right. In the first loop, ten elements of array X are copied to array Y. In the second loop, ten elements of array X are copied to array Z. Each element is one word. The starting addresses of the arrays are given below:

```
LOOP1:
For i = 0 to 9
    Y[i] = X[i];
End of For
LOOP2:
For i = 0 to 9
    Z[i] = X[i];
End of For
```

X: $0012
Y: $B010
Z: $0082

How many read misses, read hits, write misses, write hits, and block transfers occur **only** during the run of the **2nd loop** (LOOP2)? Do not count events in the first loop. Write your answers in the boxes below: (10 p)

Number of Read misses in the **2nd loop**:	0	Number of Read hits in the **2nd loop**:	10
Number of Write misses in the **2nd loop**:	1	Number of Write hits in the **2nd loop**:	9

Number of Block transfers in the **2nd loop**: 1

Student ID:
First Name and Last Name:
Signature:

QUESTION 4: (25 Points)
In a symmetric multiprocessor (SMP) system with a shared bus, there are two CPUs (CPU1 and CPU2) that have local cache memories. The system does not include a shared L2 cache.
For write operations, the **Write Back** (WB) with **Write Allocate** (WA) method is used.
There is a shared variable A in the system. To provide cache coherence, the snoopy **MESI** protocol is used.

The following questions can be answered independently.

a) Valid copies of A reside in the cache of CPU1 and in main memory (A=1). The cache memory of CPU2 is currently empty.
What is the current MESI state of the corresponding frame in the cache of CPU1? (5 p)

Exclusive

b) Valid copies of A reside in the cache of CPU1 and in main memory (A=1). The cache memory of CPU2 is currently empty. CPU2 reads the variable A.
What are the states of the corresponding frames in cache memories of CPU1 and CPU2 after the read operation? (10 p)

CPU1: Shared CPU2: Shared

c) Valid copies of A reside in the cache of CPU1 and in main memory (A=1). The cache memory of CPU2 is currently empty. CPU2 writes to the variable A (A=2).
What are the states of the corresponding frames in cache memories of CPU1 and CPU2 after the write operation? (10 p)

CPU1: Invalid CPU2: Modified