CENG3420 Homework 3

Due: Apr. 08, 2018

Please submit PDF or WORD document directly onto blackboard.
DO NOT SUBMIT COMPRESSED ZIP or TARBALL.

Q1 (**10%**) In this exercise, we will look at different ways cache capacity affects overall performance. In general, cache access time is proportional to cache capacity. Assume that main memory accesses take 68 ns. The following Table 1 data for L1 caches attached to each of two processors, P1 and P2.

Table 1: Question 1

	L1 Size	L1 Miss Rate	L1 Hit Time
P1	2 KB	13.4%	0.72 ns
P2	4 KB	7.8%	0.87 ns

1. Assuming that the L1 hit time determines the cycle time for P1 and P2, what are their respective clock rates? (5%)
2. What is the AMAT (Average Memory Access Time) for P1 and P2? (5%)

Q2 (**15%**) What are differences between interrupt and DMA? (5%) Figure 1 shows the connection among CPU, DMA control and Peripheral. Please describe the process when data is transmitted from peripheral to memory. (10%)

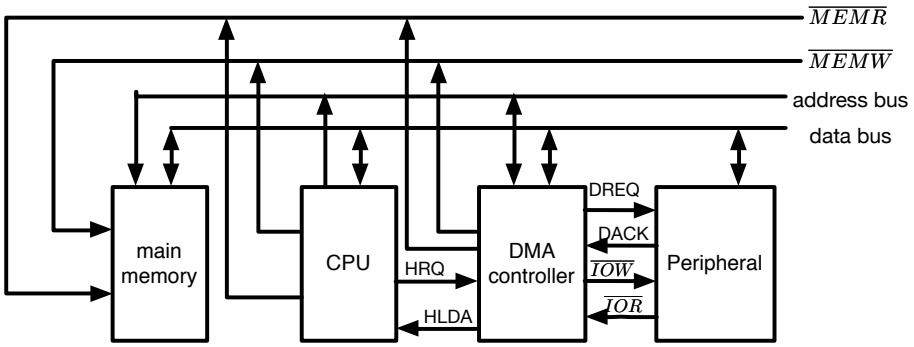

Figure 1: the connection among CPU, DMA controller and Peripheral

Q3 (**15%**) Consider the following portions of two different programs running at the same time on five processors in a symmetric multi-core processor (SMP). Assume that before this code is run, both x, y and z are 2.

```
Core 1:  x = x + 1;
Core 2:  y = z + 1;
```

```
Core 3:  w = x - y;
Core 4:  z = x + 1;
Core 5:  r = w + z;
```

1. What are all the possible resulting values of w, x, y, z and r? For each possible outcome, explain how we might arrive at those values. You will need to examine all possible interleavings of instructions. (10%)
2. How could you make the execution more deterministic so that only one set of values is possible? (5%)

Q4 (**10%**) Consider matrix multiplication $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$, where $\mathbf{C} \in \mathbb{R}^{m \times n}$, $\mathbf{A} \in \mathbb{R}^{m \times k}$ and $\mathbf{B} \in \mathbb{R}^{k \times n}$. You are given the following code to perform $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$

```
for (int i = 0; i < m; ++i) {
    for (int j = 0; j < n; ++j) {
        C[i][j] = 0;
        for (int t = 0; t < k; ++t) {
            C[i][j] += A[i][k] * B[k][j];
        }
    }
}
```

Clearly the code takes $\mathcal{O}(mnk)$ time. We would like to improve the actual running time. Please give a strategy by parallel computation without race.

Q5 (**10%**) We will upgrade a processor for CSE department. For performing application and programming, the computation speed of the new processor is 10X faster than the old processor. Assume that the old processor spends 40% time to compute and 60% to wait for I/O response. Please compute speedup.

Q6 (**15%**) In a system, the main memory size is 4 MB, the virtual memory size is 1 GB. What is the bit width of the virtual address and physical address, respectively? Assume that the page size is 4 KB, what is the **number of pages**?

Q7 (**15%**) Given an original code as follows:

```
Loop:
    L.D     F0,0 (R1)
    ADD.D   F4, F0, F2
    S.D     F4, 0 (R1)
    DADDUI  R1, R1, #-8
    BNE     R1, R2, Loop
```

1. Please revise the original code to the code with loop unrolling.(10%)
2. Based on the revised the code with loop unrolling, please revise the code with pipeline scheduling.(5%)

Q8 (**10%**) We will examine space/time optimizations for page tables. Table 2 shows parameters of a virtual memory system.

Table 2: Question 8

	Virtual Address (bits)	Physical DRAM Installed	Page Size	PTE Size (byte)
a	43	16 GB	4 KB	4
b	38	8 GB	16 KB	4

1. For a single-level page table, how many page table entries (PTEs) are needed? How much physical memory is needed for storing the page table? (5%)
2. Using a multilevel page table can reduce the physical memory consumption of page tables, by only keeping active PTEs in physical memory. How many levels of page tables will be needed in this case? And how many memory references are needed for address translation if missing in TLB? (5%)