Homework #4

Muhammed Orhun GALE - 26754

Assigned: 25/05/2023
Due: 08/06/2023

1. (**15 pts**) Consider a processor with the following parameters:

Base CPI, no stalls due to cache misses	0.5
Processor speed	2.5 GHz
Main memory access time	100 ns
First-level cache miss rate per instruction	3%
1st option for the second-level cache	
Direct-mapped: the latency	20 cycles
Local miss rate of the second level cache, direct mapped	50%
2nd option for the second-level cache	
8-way set associative: the latency	50 cycles
Local miss rate of the second level cache, 8-way set associative	40%

   a. (**5 pts**) We are considering two alternatives for second-level cache memory as shown above. What are the global miss rates if we use second-level direct-mapped cache (1st option) and second-level 8-way set-associative cache (2nd option)?

      Global miss rate if we use 1st option: 3% * 50% = 1.5%

      Global miss rate if we use 2nd option: 3% * 40% = 1.2%

   b. (**10 pts**) Calculate the CPI for the processor in the table using: i) only first-level cache, ii) a second-level direct mapped cache, and iii) a second-level 8-way set associative cache.

(i) First find how many clock cycles does it take to access main memory —> 10^-7 seconds/(1/(2.5 * 10^9 seconds)) —> 250 clock cycles

Then CPI = 0.5 + (3% * 250) = 8

(ii) CPI = 0.5 + (3% * 20) + (1.5% * 250) = 0.5 + 0.6 + 3.75 = 4.85
(iii) CPI = 0.5 + (3% * 50) + (1.2% * 250) = 0.5 + 1.5 + 3 = 5

2. (**15 pts**) Assume a direct-mapped cache with 16-byte cache lines. The following code is written in C programming language, where elements of integer arrays of **A** and **B** within the same row are stored contiguously in memory.

```
for(i=0; i<8; i++)
    for(j=0; j<8; j++)
        A[i][j] = A[i][j] + B[i]
```

Assuming there is no conflict and capacity misses, compute the miss rates for this code due to compulsory cache misses.

Since each cache line is 16-byte, a line can store 4 integers so while accessing for 1 integer, next 3 elements can be cached too.

On first iteration —> Miss for A[0][0] —> A[0][1-2-3] cached and miss for B[0] —> B[1-2-3] cached

Then, every 4th iteration is a miss —> 1-2-3 / 5-6-7 / …. / 61-62-63

Therefore —> 64/4 = 16 miss for A + 2 miss for B —> 18 miss in total
Total cache access —> 2 times for A and B on each iteration —> 128

Then —> miss rate due to compulsory cache misses = 18/128

3. (**12 pts**) Consider Intrinsity FastMath Processor that implements MIPS 32 instruction set architecture. Its virtual addresses are 32-bit integers. It uses 16 KB pages. The processor has a cache with 256 entries where the block size is 16 words (64 B). Assume that a virtual address is 32-bit number, shown as Address[31:0]. Adopting the notation used for byte offset, Address[1:0] for instance, answer the following questions.

   a. (**3 pts**) Give the address bits, namely Address[?:?], for cache index.

2^8 cache lines —> 4-bit cache index + 4-bit block offset

—> Address[9:6]

   b. (**3 pts**) Give the address bits, namely Address[?:?], for page offset.

2^14 lines

—> Address[13:0]

   c. (**3 pts**) Give the address bits, namely Address[?:?], for virtual page number.

—> Address[31:14]

   d. (**3 pts**) Give the address bits, namely Address[?:?], for block offset.

—> Address[6:2]

4. (**15 pts**) Consider a hard disk with the following parameters:

   - Rotation speed: 7200 RPM
   - Average seek time: 5.8 ms
   - Transfer rate: 2 MB/s
   - Controller overhead: 8 ms
   - Sector size: 4096 B

   a. **(7 pts)**What is the average time to read or write a sector from a disk?

One rotation —> 60/7200 = 0.0083 —> 8.33/2 = 4.17 ms —> average rotational latency
Data transfer —> 4096/(2^20) = 0.00391 = 3.91 ms

Average time to read = 5.8ms + 4.17ms + 3.91ms + 8 = 21.88 ms

   b. **(8 pts)** If the transfer rate increases to 4 MB/sec and the control overhead decreases to 6 ms, how much faster is the new disk system?

New transfer time = 4096/(4 * 2^20) = 0.00195
Average time to read = 5.8ms + 4.17ms + 1.95ms + 8 = 17.92 ms
Speed-up = 21.88ms - 17.92ms = 3.96ms

5. Answer the following questions (**15 pts**)

a. (**5 pts**) Assume that you have a hard disk with MTTF = 400,000 hours/failure. Calculate the annual failure rate of the disk (AFR).

AFR = 1 - exp(-8760/400000)
1 - exp(-0.0219) == 0.0216 —> 2.16%

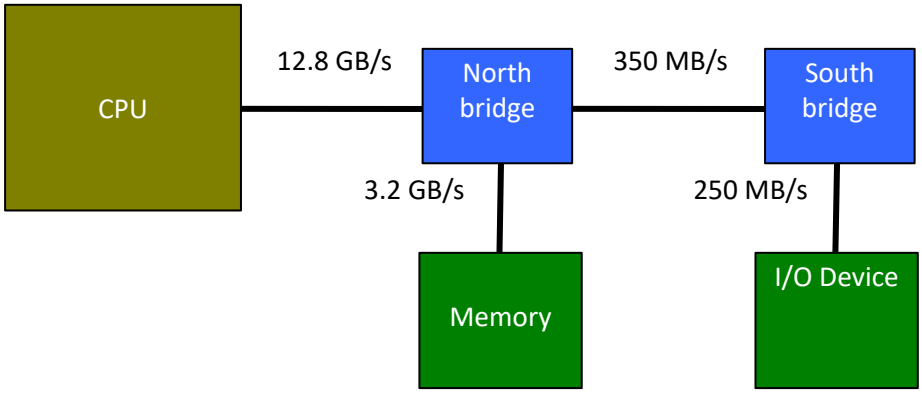
b. (**5 pts**) You are asked to construct a **RAID1** disk system using two disks with capacities 100 GB and 150 GB, respectively. Find the total usable size.

Usable size == smallest disk's capacity —> 100GB (since duplicated)

c. (**5 pts**) Disk failures are not independent and if one disk fails the other fails with 50% probability, as well. What is the overall failure rate of **RAID1** you constructed in (b) using the AFR you find in (a)?

2.16% * 0.5 = 1.08%

6. (**15 pts**) Consider the computer system shown in the figure below:



The bandwidths of the devices are:

Device	**Bandwidth (Byte/s)**
Front Side Bus	12.8×10^9
Memory	3.2×10^9
North-South Bus	350×10^6
I/O Device	250×10^6

Assume that the CPU needs an average of 4 bits from memory and 0.2 bits from the I/O device in order to execute an instruction while running a program.

a. (**10 pts**) Show the maximum instruction execution rates (in MIPS) that each device can sustain. What is the maximum sustainable MIPS of the system (assume that the CPU is not the bottleneck in the system)?

Front Side Bus:

BW = 12.8 * 10^9 * 8 bits/s

CPU requires 4+0.2 bits —> (12.8 * 10^9 * 8 bits/s) * 1/4.2 = 24.3 * 10^9 inst/s —> 24300 MIPS

Memory:

BW = 3.2 * 10^9 * 8 bits/s

CPU requires 4 bits —> (3.2 * 10^9 * 8) * 1/4 = 6.4 * 10^9 inst/s —> 6400 MIPS

North/South Bridge Bus:

BW = 350 * 10^6 * 8 bits/s

CPU requires 0.2 bits —> (350 * 10^6 * 8 bits/s) * 5 = 5.6 * 10^9 inst/s —> 14000 MIPS

I/O device:

BW = 250*10^6* bits/s

CPU requires 0.2 bits —> (250*10^6* 8 bits/s) * 5 = 10^10 inst/s —> 10000 MIPS

Sustainable MIPS is 6400 since the bottleneck is 6400 with Memory

b. (**5 pts**) Suppose that the CPU has a clock speed of 1.8 GHz and it can execute 3 instructions per cycle on average (since it is a multi issue processor). What is the bottleneck in the system now?

CPI = 0.33
MIPS = (1.8 GHz / 0.33) / 10^6 = 5400 MIPS

Now CPU becomes the bottleneck since 5400 MIPS is the smaller than 6000 MIPS with memory.

7. (**15 pts**) Assume that a program executes in 100 seconds, where 80 seconds is CPU time and the rest of the time is spent for I/O operations. If the CPU performance improves by a speedup value of 1.5 and I/O performance improves by a speedup value of 1.1 per year, how much faster will the program run after 5 years?

Year 0 -> CPU Time =80 s - I/O Time 20 s - Elapsed Time 100 s
Year 1 -> CPU Time =53.33 s - I/O Time 18.18 s - Elapsed Time 71.51 s
Year 2 -> CPU Time =35.55 s - I/O Time 16.52 s - Elapsed Time 52.07 s
Year 3 -> CPU Time =23.7 s- I/O Time 15.02 s - Elapsed Time 38.72 s
Year 4 -> CPU Time =15.8 s - I/O Time 13.65 s - Elapsed Time 29.45 s
Year 5 -> CPU Time =10.53 s - I/O Time 12.41 s - Elapsed Time 22.94 s

Speed-up = 100s /24s = 4.36