CDT204 – Computer Architecture

Date: Nov. 8th 2018

Time: 8:10 – 12:30

Help: Small calculator is allowed in the exam

The exam has 39 points. The grades will be awarded as follows:

- 3 : 20 Points
- 4 : 28 Points
- 5 : 36 Points

Important Notes:

- Give as full an answer as possible to obtain full marks. All calculations, approximations, assumptions and justifications must be reported for full credit unless stated otherwise. Please use figures and examples to clarify.
- If you do not understand a question clearly, make an assumption, write the assumption in the paper and solve according to that.
- Write the question and part number on each page clearly.
- Answer each question on a separate page.
- Do not use pencil. NO PENCIL.
- In case you might have forgotten:

  $1GB = 2^{10}MB = 2^{20}KB = 2^{30}B$

  $1\ sec = 10^{3}\ ms = 10^{6}\ \mu s = 10^{9}\ ns = 10^{12}\ ps$

Live long and prosper

Task 1 – General (4p)

Signed integer numbers are stored in the computer memory using "Two's Complement" method.

1. Explain this method with examples. (2p)
2. Why is "two's complement" the most widely used method for storing signed integers? (2p)

Task 2 – Performance (10p)

Dr. R. Chandra has a computer with the following properties:

	Property	Value
Cache System	Instruction Cache Miss Rate	2%
	Data Cache Miss Rate	5%
	Cache Miss Penalty	400 cycles
Instruction Class CPIs (ideal system)	Load/Store*	3
	Branch	2
	Arithmetic	1

**Cache/Memory penalties not included.*

For the type of work that Dr. Chandra uses the computer, the average distribution of different instruction classes (types) is as below:

	Load/Store	Branch	Arithmetic
Percentage	30%	20%	50%

Recently, Dr. Chandra has received some funding and he can use part of the budget to upgrade the computer. Unfortunately, the budget is not enough to upgrade everything and he has to pick one of the following options:

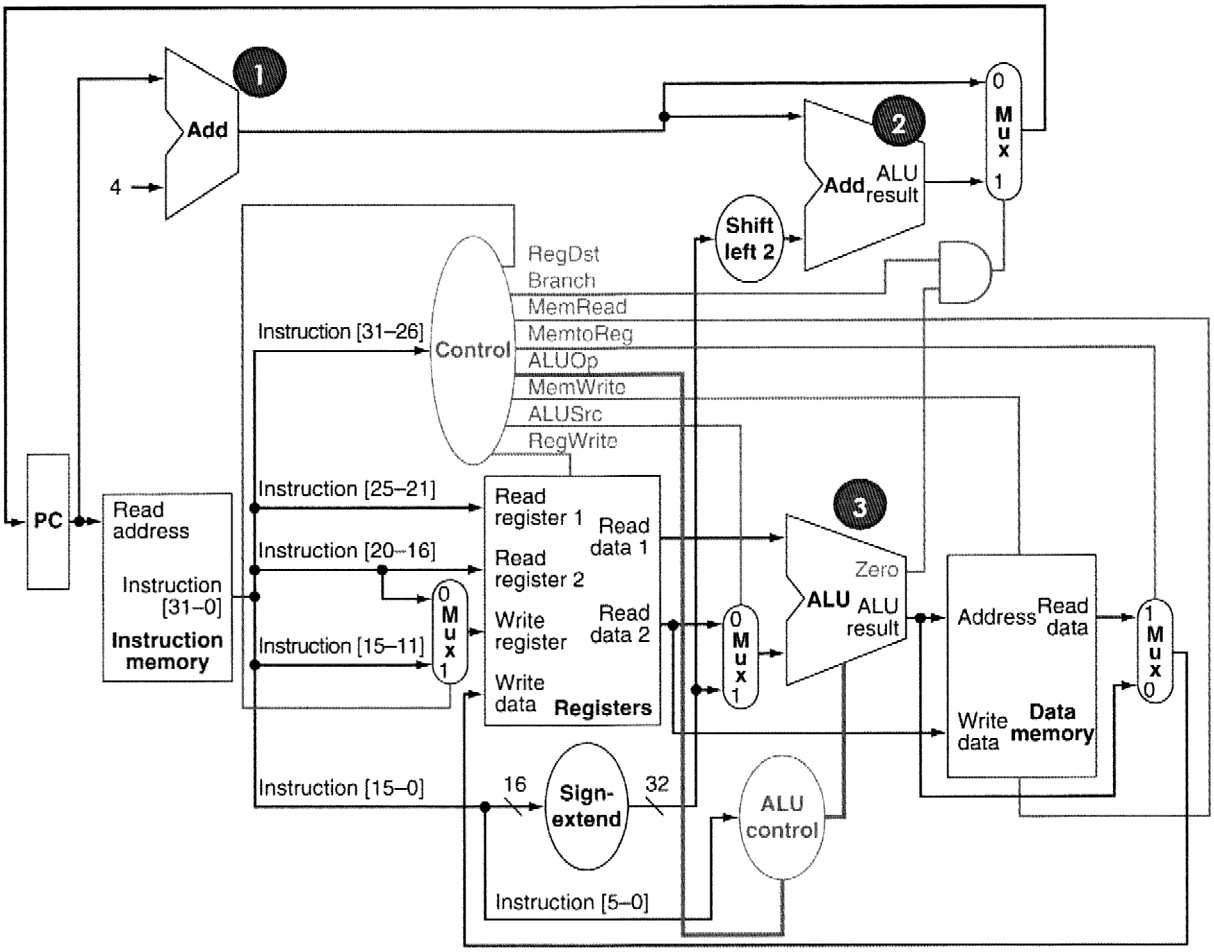
1. Purchase a machine with a processor and cache that are twice as fast as the current system (memory system has the same speed as current device).
2. Purchase a machine with a processor and cache that are the same speed as current machine but in which the cache is twice as large. Assume that the cache miss rate will drop by 40% on the larger cache (although this is generally not true in real life).
3. Purchase a machine with a processor and cache that are the same speed as current machine but also include an L2 cache with 1% instruction miss rate and 2% data miss rate (these are not the global miss rates), assume that L2 cache has a latency of 20 cycles.

Since Dr. Chandra is busy developing his doomsday device, he has asked me to do the calculations and choose the best possible option for him. I, in the other hand, have decided to leave it to my students to do that. So please! Enlighten us! what should we do? Which upgrade should we purchase? How much faster will the chosen upgrade be?

(Grading: Performance of the original system: 3p, upgrades: 2p each, relative performances: 1p)

Task 3 – Datapath (8p)

The following diagram shows a simple datapath for MIPS instruction set. The datapath contains 3 ALUs marked by number 1-3.



A. Explain the job of each of these ALUs in detail. Use examples. (6p)

B. The output of ALUs number 1 and 2 goes to a Mux. What is the purpose of this Mux? What is the other input to it? (2p)

Task 4 – Pipelining (3p)

Name and explain different types of hazards that might occur in a pipelined datapath using examples. (3p)

Task 5 – Virtual Memory (6p)

Suppose that you have a page size of 4 KB. Also suppose that you have an application in which the text segment (program code) takes up 2 KB, the data segment (global variables, etc.) takes up 5 KB, and that 1 KB allocated for the stack.

A. We do not know where in the process's private memory space these three segments are allocated, but we know that they are in 3 distinct address ranges. What is the maximum number of pages that this program might need? (1p)

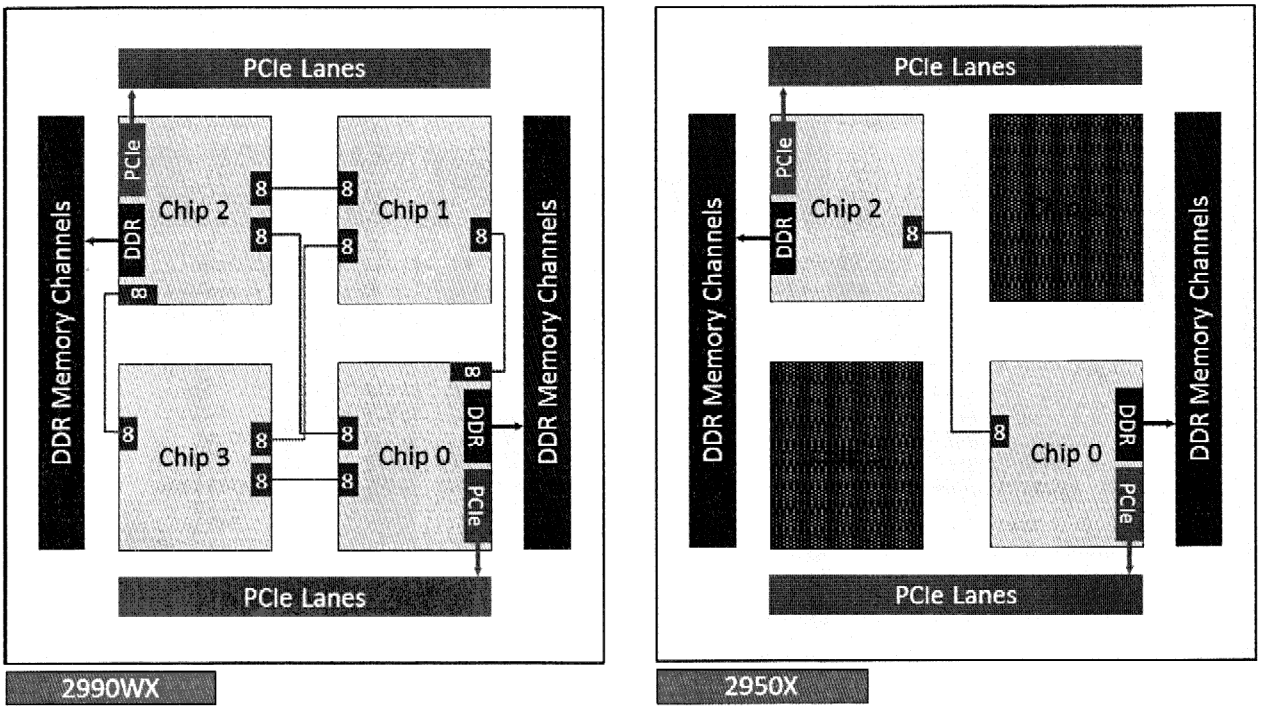
Now assume that we know that the code segment is allocated in address 1000_{16} to $17FF_{16}$, data segment from 10000_{16} to $113FF_{16}$ and the stack from $FFC00_{16}$ to $FFFFF_{16}$ in a 1MB virtual memory. After the program runs for a while, some parts of the page table look like the following:

Valid	Dirty	Virtual Page No.	Physical Page No.
1	0	00010000	001100101
1	0	00000001	000110011
1	1	11111111	110110100
0	0	00000001	111001000
1	1	00010001	000110101

B. Considering that the virtual memory is 1MB in size, how many bits the virtual addresses should be? (1p)
C. Judging by the physical page numbers, how big is the physical memory? (1p)
D. What page number the code segment has got? (The whole segment is placed in one page) (2p)
E. Is all of the data segment read into the physical memory? (2p)

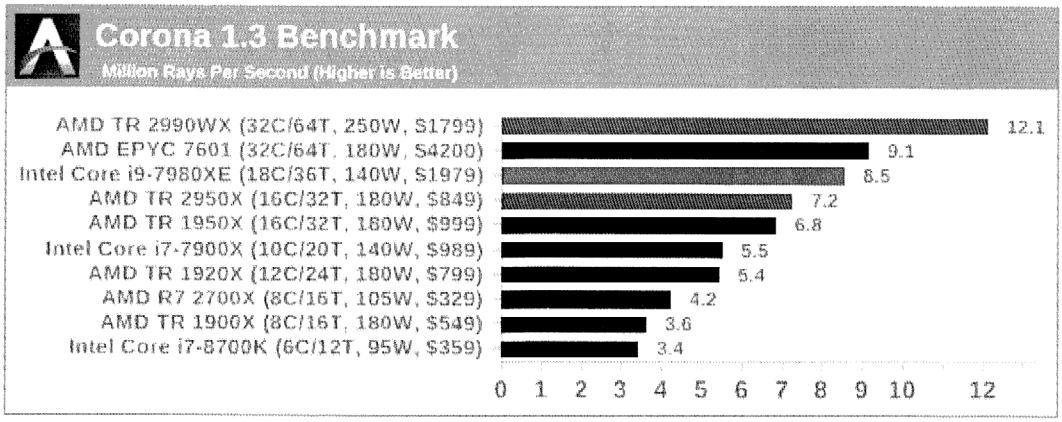
Task 6 – Multiprocessing (8p)

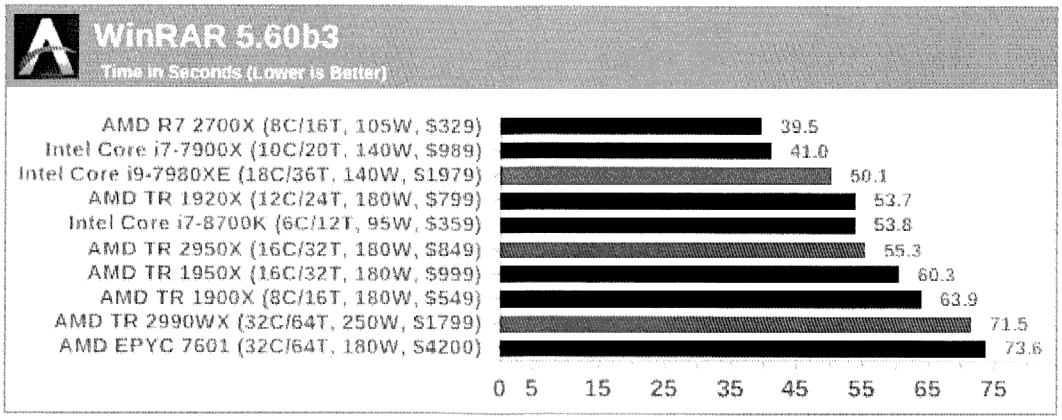
In recent years, AMD has changed the landscape of High End DeskTop (HEDT) market for the CPUs. Two of their latest processors in HEDT segment of the market are Threadripper 2950X and Threadripper 2990WX. The 2950X part, offers 16 cores and 32 threads while the 2990WX part offers 32 cores and 64 threads. Both of these CPUs are based on a smaller chip which contains 8 cores/16 threads and both of these CPUs contain 4 of these smaller chips. The schema bellow shows the overall view of the 2990WX and 2950X designs.



In 2990WX four of the smaller 8 cores chips are connected to each other through special connections called infinity fabric (the red line with the infinity sign on them [those are infinity not number eight]), making it a CPU with a total of 4 x 8 = 32 cores. The 2950X uses the same design (4 smaller chips), but two of the smaller chips are disabled (marked with dark color) giving it 2 x 8 = 16 cores. In both designs if one of the smaller chips require a resource which is available in another one of the smaller chips, they need to access it through infinity fabric.

Reviewing website Anandtech, ran some benchmarks on these CPUs. Here two of those results are presented with the above Threadripper CPUs shown in orange (next page):





A. In Corona benchmark the 2990WX performs 68% better than the 2950X. What does this result indicate about the benchmark itself? (2p)

B. In WinRAR benchmark, interestingly, the 2950X performs almost 30% faster than 2990WX although it has half the number of cores. Why? What does this result indicate about the benchmark itself? (2p)

C. What would be your guess about the result, if we run a gaming benchmark on these processors? Which one of them might perform better? Why? (2p)

D. Which one of these processors can achieve a higher clock speed theoretically? Please discuss. (2p)