

CSE141 HOMEWORK 1

UC San Diego
WINTER 2014

ISA

Note: Some problems say SB which means to see book as the problem references a chart or figure from the book

2.4.1. For the C statement $[f = g - A[B[4]]];$

- What is the corresponding MIPS assembly code?
- How many registers are needed?

2.18.5a (SB) Translate the loops into C. Assume that the C-level integer i is held in register $\$t1$, $\$s2$ holds the C-level integer called result, and $\$s0$ holds the base address of the integer MemArray.

P1. By how many bits should the immediate field of an I-type instruction be reduced to match the length on an R-type instruction for the follow modifications to the MIPS architecture?

- [ISA-A] 8 general purpose registers.
- [ISA-B] 128 general purpose registers.
- Given ISA-A and ISA-B describe programs that are more efficiently encoded with one rather than the other.

P2. The instruction subtract immediate [Usage: $\text{subi } \$t2, \$t3, 5$] does not exist in MIPS.

- What would be an alternate instruction that would work in almost all the same cases?
- There is one case where it will not work, what is it?
(Hint: What range of values can be given in the immediate field)

Performance

1.15.1a (SB) How much is the total time reduced if the time for FP operations is reduced by 20%?

1.16.1b (SB) Find the total execution time and by how much it is reduced if the time of routines A, C, and E is improved by 15%.

1.16.2b (SB) How much is the total time reduced if routine B is improved by 10%?

1.16.3b (SB) How much is the total time reduced if routine D is improved by 10%?

P1. Define and compare throughput and latency in your own words. Give examples where throughput increases at the cost of latency and vice versa (these examples do NOT have to be computer architecture related).

P2. Decide between building a processor which executes at 1.5GHz and has an average CPI of 2.2 and a processor which executes at 2GHz, but has a CPI of 2. Which is better to build and why?

P3: Assume that you propose an optimization in which the memory latencies are sped up by a factor of 4.4. After implementation, you observe that you now spend half your time on waiting for memory. What percentage of the original execution, i.e. before the optimization, was spent waiting for memory?

P4. The release of Intel's Sandy Bridge processors was unique in its integration of the GPU and CPU on the same die. Compare the performance of having the CPU and GPU in separate packages versus the same die.

System A: CPU and GPU are on the same die. The GPU gives a 7 times speedup for the 95% of the program.

System B: CPU and a discrete GPU in separate packages. In this case the GPU is able to give a speedup to 75% of the program.

How much of a speedup must the discrete GPU provide in order for the overall performance to equal that of the integrated system?