

Lecture 3:

Exercise 9:

1. Let p denote performance, and I the instruction count (multiplied by some constant so we can use the GHz unit). Then $I * p_x = (\text{clock rate of } P_x) / (\text{CPI of } P_x)$, and so we can compare the three processors; $I * p_1 = 2/1.5 = 40/30$, $I * p_2 = 1.5/1 = 45/30$, $I * p_3 = 3/2.5 = 36/30$. Meaning, compared between the three, P2 has the highest performance.
2. We get these by juggling around the various expressions for performance terms. To be brief, in order, number of cycles: $2 * 10^{10}$, $(3/2) * 10^{10}$, $3 * 10^{10}$, and number of instructions: $(4/3) * 10^{10}$, $(3/2) * 10^{10}$, $(6/5) * 10^{10}$.
3. By manipulating the fact that $0.7 * (\text{CPU time}) = (\text{Instructions}) * 1.2 * (\text{CPI}) / (\text{New clock rate})$ we get that $(\text{New clock rate}) = (12/7) * (\text{Instructions}) * (\text{CPI}) / (\text{CPU time}) = (12/7) * (\text{Old clock rate})$. So to reduce the CPU time by 30% we would need to increase the clock rate by roughly 71%, so, in order, P1, P2, P3 would need to become 3.4 GHz, 2.6 GHz and 5.1 GHz, respectively.

Exercise 10:

1. a. $200 - 0.2 * 35 = 193\text{s}$ b. $210 - 0.2 * 50 = 200\text{s}$ 2. a. $0.2 * 85 = 17\text{s}$, b. $0.2 * 80 = 16\text{s}$ 3. No, even if branch instructions were eliminated completely there would only be time reductions of 15% and 14% for a and b respectively.

Exercise 11:

1. Compute time ratios: $[0.55, 0.51, 0.61, 0.47, 0.48]$, Communication time ratios: $[1.18, 1.31, 1.29, 1.05, 1.13]$.
2. Geometric means: 0.52 and 1.19. So we would suspect getting a computation time of $0.52 * 6.5 = 3.38\text{ms}$ and the time for communication to be $26 * 1.19 = 30.94\text{ms}$.
3. By the same extrapolation as above we can approximate the times for a single processor as $176/0.52 = 338.46\text{ms}$ for computation and $11/1.19 = 9.24\text{ms}$ for communication.

Exercise 12:

1. *Ideally*; 2500, 1250, 625 requests/sec for 4, 8, 16 CPUs/cores respectively.
2. This would mean that *everything* performed in a 'request' is completely parallelizable, which is quite unlikely.