CSCE 692 - Chapter 1 Homework Problems

(100 Points)

Due: NLT 1100 Tuesday, 15 January 2019

**Problems: 1.1, 1.2, 1.4, 1.7, 1.8, 1.16 (5 pts / subproblem)**

**Instructions:**

* Print your name on each page
* Clearly indicate your answer
* Clearly show your work such that I can understand your thinking
* Explain any assumptions and provide references (sources) for any additional information you had to research in order to complete any problem.

Recommendations (i.e., how to score max points)

* Read through the problem statement in this handout, which includes important clarifications from the text.
* Think through and work out the problems on scratch paper first, then copy your answers (electronically, or by hand) into this handout.

**Case Study 1: Chip Fabrication Cost** 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Chip** | **Die Size (mm2)** | **Estimated defect rate**  **(per cm2)** | **N** | **Manufacturing size (nm)** | **Transistors (billion)** | **Cores** |
| BlueDragon | 180 | 0.03 | 12 | 10 | 7.5 | 4 |
| RedDragon | 120 | 0.04 | 14 | 7 | 7.5 | 4 |
| Phoenix8 | 200 | 0.04 | 14 | 7 | 12 | 8 |

**Figure 1.26 Manufacturing cost factors for several hypothetical current and future processors**

1. [5/5] <1.6> Figure 1.26 gives the hypothetical relevant chip statistics that influence the cost of several current chips. In the next few exercises, you will be exploring the effect of different possible design decisions for the Intel chips.
   1. [8] <1.6> What is the yield, or percentage of good dies, for the Phoenix chip?
   2. [5] <1.6> Why does the Phoenix chip have a higher defect rate than BlueDragon?

The Phoenix chip has a higher defect rate because . N is a measure of the manufacturing complexity of a chip; thus, the Phoenix chip is harder to produce giving it a higher defect rate.

1. [5/5] <1.6> They will sell a range of chips from that factory, and they need to decide how much capacity to dedicate to each chip. Imagine that they will sell two chips. Phoenix is a completely new architecture designed with 7 nm technology in mind, whereas RedDragon is the same architecture as their 10 nm BlueDragon. Imagine that RedDragon will make a profit of $15 per defect-free chip. Phoenix will make a profit of $30 per defect-free chip. Each wafer has a 450 mm (45 cm) diameter.
   1. [5] <1.6> How much profit do they make on each wafer of Phoenix chips?

, after rounding down the dies per wafer to the nearest integer.

* 1. [5] <1.6> How much profit do they make on each wafer of RedDragon chips?

Profit = $9602.50

**Case Study 2: Power Consumption in Computer Systems**

1.4 [5/5/5/5] <1.5> A cell phone performs very different tasks, including streaming music, streaming video, and reading email. These tasks perform very different computing tasks. Battery life and overheating are two common problems for cell phones, so reducing power and energy consumption are critical. In this problem, we consider what to do when the user is not using the phone to its full computing capacity. For these problems, we will evaluate an unrealistic scenario in which the cell phone has no specialized processing units. Instead, it has a quad-core, general-purpose processing unit. Each core uses 0.5 W at full use. For email-related tasks, the quad-core is 8× as fast as necessary.

1. [5] <1.5> How much dynamic energy and power are required compared to running at full power? First, suppose that the quad-core operates for 1/8 of the time and is idle for the rest of the time. That is, the clock is disabled for 7/8 of the time, with no leakage occurring during that time. Compare total dynamic energy as well as dynamic power while the core is running.\

because it is completing the same tasks.

The Dynamic power is reduced to 1/8 of the previous power because the same tasks were accomplished in 1/8 of the time.

1. [5] <1.5> How much dynamic energy and power are required using frequency and voltage scaling? Assume frequency and voltage are both reduced to 1/8 the entire time.

1. [5] <1.6, 1.9> Now assume the voltage may not decrease below 50% of the original voltage. This voltage is referred to as the voltage floor, and any voltage lower than that will lose the state. Therefore, while the frequency can keep decreasing, the voltage cannot. What are the dynamic energy and power savings in this case?

1. [5] <1.5> How much energy is used with a dark silicon approach? This involves creating specialized ASIC hardware for each major task and power gating those elements when not in use. Only one general-purpose core would be provided, and the rest of the chip would be filled with specialized units. For email, the one core would operate for 25% the time and be turned completely off with power gating for the other 75% of the time. During the other 75% of the time, a specialized ASIC unit that requires 20% of the energy of a core would be running.

**Exercises**

1.7 [5/5/5/5/5] <1.4, 1.5> One challenge for architects is that the design created today will require several years of implementation, verification, and testing before appearing on the market. This means that the architect must project what the technology will be like several years in advance. Sometimes, this is difficult to do.

1. [5] <1.4> According to the trend in device scaling historically observed by Moore’s Law, the number of transistors on a chip in 2025 should be how many times the number in 2015?

Moore’s law = double every 2 years

1. [5] <1.5> The increase in performance once mirrored this trend. Had performance continued to climb at the same rate as in the 1990s, approximately what performance would chips have over the VAX-11/780 in 2025?

In 2003, performance was 6,043 times faster than the VAX-11/780. From that point, if performance continued to increase at 52%/year, 22 years later in 2025 performance would be:

times faster

1. [5] <1.5> At the current rate of increase of the mid-2000s, what is a more updated projection of performance in 2025?

Current Rate = 3.5%/year, 49,935 times faster in 2017.

times faster

1. [5] <1.4> What has limited the rate of growth of the clock rate, and what are architects doing with the extra transistors now to increase performance?

Power restricts the growth of clock rate. We are doing the following things to increase performance:

* Improving race to halt 🡪 saving power by stopping work faster
* Do nothing well
* Dynamic Voltage and Frequency Scaling (DVFS)
* Designing for the typical case
* Overclocking

1. [5] <1.4> The rate of growth for DRAM capacity has also slowed down. For 20 years, DRAM capacity improved by 60% each year. If 8 Gbit DRAM was first available in 2015, and 16 Gbit is not available until 2019, what is the current DRAM growth rate?

8 Gbit in 2015, 16 Gbit in 2019

The current growth rate is 18.9%/year.

1.8 [5/5] <1.5> You are designing a system for a real-time application in which specific deadlines must be met. Finishing the computation faster gains nothing. You find that your system can execute the necessary code, in the worst case, twice as fast as necessary.

1. [5] <1.5> How much energy do you save if you execute at the current speed and turn off the system when the computation is complete?

1. [5] <1.5> How much energy do you save if you set the voltage and frequency to be half as much?

1.16 [5/5/5/5/5] <1.10> When parallelizing an application, the ideal speedup is speeding up by the number of processors. This is limited by two things: percentage of the application that can be parallelized and the cost of communication. Amdahl’s Law takes into account the former but not the latter.

1. [5] <1.10> What is the speedup with N processors if 80% of the application is parallelizable, ignoring the cost of communication?
2. [5] <1.10> What is the speedup with eight processors if, for every processor added, the communication overhead is 0.5% of the original execution time.
3. [5] <1.10> What is the speedup with eight processors if, for every time the number of processors is doubled, the communication overhead is increased by 0.5% of the original execution time?
4. [5] <1.10> What is the speedup with N processors if, for every time the number of processors is doubled, the communication overhead is increased by 0.5% of the original execution time?
5. [5] <1.10> Write the general equation that solves this question: What is the number of processors with the highest speedup in an application in which P% of the original execution time is parallelizable, and, for every time the number of processors is doubled, the communication is increased by 0.5% of the original execution time?