CS 6810 Homework #1

Due: 9:10 a.m. Sept. 8 (no late submissions will be graded)

General instructions: You will hand in your homework's at the beginning of class in paper form. That way the TA can make comments during the grading process. Note that if we can't read your writing then you will be penalized. Similarly, if you just write the answer without supporting it with your reasoning, you will be penalized. Hence you are encouraged to type your answers where this makes sense. Note this is often not easy when doing algebraic calculations unless you have a good function editor, or use MatLab or whatever. The key is that your solutions must be readable in order to receive credit. Organize your work accordingly. The general intent is to make the problem statement clear but there will be times when this goal unfortunately isn't met. Clarifications can be sought by email but if you're in a middle of the night scramble the chances of an instant answer are low. If you make reasonable assumptions and clearly state them then you won't be penalized. "Googling for answers" is unfortunately a common practice, and given the broad use of your textbook in other universities it is not unlikely that you will find problems and answers similar to these somewhere. If you rely on this mechanism you will likely lose when it comes to exam time. Make sure that what you turn in is your own work – see the cheating policy on the class web page. However you are encouraged to discuss solution strategies with your classmates, just make sure that you solve the problems on your own.

Problem 1: [35 points] HoHum, Mini, Maxi, and Moby

In this problem you will use the data in Figure 1.22 in your text. You always have too many choices and you usually need to decide based on quantitative estimates. The point of this exercise is to illustrate some of these issues.

Also assume that the wafer diameter in the new 90 nm process is 300 mm, and that you will be able to sell everything you make. Your fab capacity is 5,000 wafers per month, and you predict that this production run will last 18 months. Also assume that the α and defect rate doesn't change during this time. [Note that this is not in general true due to the learning curve issues discussed in class, but it does simplify the problem.]

1.1 [1 point] Why does the Power5 have a lower defect rate than the Niagra and Opteron?

1.2 [5 points] What is expected die-yield for the Power 5?

1.3 [29 points] 4 design options – which one is the best?

The new 90nm process is available and you are trying to decide what enhancement of the Power 5 will become the new Power 6 – each option has a different internal name. Assume that the area of an existing design (codename

HoHum in this exercise) will be 5% larger than the idealized prediction of the process shrink. Similarly assume that the processor performance enhancement of HoHum due to the shrink is 20% worse than the idealized performance due to faster transistors.

a) **[2 points]**
Why is the die area larger than the idealized prediction?

b) **[2 points]**
Why doesn't performance track transistor speed up in the new process?

c) **[25 points]**
In order to simplify the problem, all of the usual costs of validation, testing, packaging, power, pin-cost, etc. are ignored and you only need to consider the economic argument using a simple expected profit value. The 4 designs are as follows:

HoHum – just take the existing 130nm Power 5 and shrink it for 90nm. Expected profit for this option is $50 per device.

Mini – Enhance the arithmetic units (both FP and Int). These enhancements have been calculated to increase the die area of HoHum by 4% and increase performance over HoHum by 7%. Expected profit is $55 per processor.

Maxi – use the Mini enhancements, and increase the cache sizes.. This will increase the area over HoHum by 18% and increase the performance by 19%. Expected profit per processor is $65.

Moby – do everything that Maxi does but make the processor dual core and multi-threaded. This will increase the area over HoHum by 45% but double the performance over HoHum. Expected profit per processor is $115.

Management only cares about profit – which option do they choose? Clearly show your calculations and reasoning for your choice!!

Problem 2: [25 Points] Power, Performance, and Cost

In today's microprocessor world, power is the 1st order constraint since it has a very direct effect on performance and cost. The point of this exercise is to illustrate some of the tradeoff's. Several unrealistic assumptions will be made in this problem in order to simplify things but in reality this is an extremely complex set of issues.

Once again we have 4 designs that we will consider (A, B, C, & D – no hokey names this time). Cost in this case is a packaged working part price. For this exercise we will *unrealistically* ignore issues such as reliability, testing cost, and yield. An important thing to understand is that half of the power consumed in a processor is given off as heat and the rest is returned to the power supply. All of today's processors protect themselves from damage due to excessive heat and for this exercise we will assume that this protection mechanism involves sensors and the ability of a temperature monitor to reduce clock frequency. However reducing clock frequency is not simple and changing the clock frequency requires the processor to be halted until the various timing elements stabilize (PLL's, DLL's, etc.). In general the stabilization delay will vary with process, design, temperature, and target clock frequency. In this exercise we will assume the stabilization delay is always 1 millisecond. When a processor gets to it's thermal threshold (which depends on packaging and the cooling capability of the system) it will finish the things that are currently in the pipeline, halt, switch frequency to the next lower setting, stabilize, and then resume processing.

Note that for this exercise you will choose which processor provides the best performance per dollar. This is clearly an oversimplification since what you buy is a system where the processor is just one or more of the components. In this exercise we will focus solely on the processor.

The following table contains data for the 4 processors to be studied – each one can run at 4 frequencies. The % values, next to the frequency in GHz, indicates the percentage of the unstalled time that the processor spends operating at this frequency. The performance is based on SPEC2006 and is normalized to the slowest design. This peak performance is based on simulation data that assumes that the processor is always running at Freq1. However, after thermal analysis, the actual time that the processor spends at a given frequency is calculated.

Processor	Peak Perf (at Freq1)	Cost	Freq1 (GHz)	Freq2 (GHz)	Freq3 (GHz)	Freq 4 (GHz)	% stall
A	1	$250	2.1/80%	1.8/15%	1.5/5%	1.1/0%	2%
B	1.2	$400	2.7/75%	2.2/10%	1.8/10%	1.1/5%	4%
C	1.8	$600	2.9/80%	2.2/20%	1.8/0%	1.1/0%	3%
D	2.7	$900	3.1/70%	2.8/10%	2.5/10%	2.0/10%	6%

1.1 [3 points]

[Note: there is no exact right answer for this one – it is just a way for use to get a handle on your architectural savvy at this early point in the course – you will be graded based on making reasonable arguments.]

Provide an explanation of what architectural differences could be present to justify the performance and frequency values.

1.2 [2 points]

Assume you were told that processor B and C were exactly the same.

What is wrong with the data in the table?

Why is C 50% more expensive than B?

1.3 [3 points]

Prior to thermal analysis which processor has the best performance/dollar?

1.4 [15 points]

After thermal analysis, which processor has the best performance/dollar?

1.5 [2 points]

If you are buying a new machine and have looked at the SPEC ratings of the processor, what else do you need to consider to get the most from your dollars?

Problem 3: [15 points]: Performance

You have a potential new processor design which improves the CPI of the baseline processor, but the cycle time of the new innovation increases by 5% due to the additional complexity of the additional circuitry. Assume *unrealistically* that all Instructions take the same amount of time in this exercise.

Given the following table:

	BM1	BM2	BM3	BM4	BM5
# instructions	1.6M	4.1M	3.2M	2.1M	2.2M
Baseline IPC	1.8	2.3	1.6	2	0.7
Innovation IPC	2.1	2.6	1.7	2.9	0.9

BMn denotes a particular benchmark which when fully executed has executed a certain number of instructions indicated in the 2nd row. The baseline IPC =1/CPI is given and the new IPC with your innovation is shown in the 3rd row.

1.1 [3 points]

How much better is your new design based on an arithmetic mean?

1.2 [3 points]

How much better is your new design based on a geometric mean?

1.3 [3 points]

Which of the above 2 values do you report to your marketing department and why.

1.4 [3 points]
How much better is your new design for a client with the same IP profile for a set of benchmarks, each of which will run for the same amount of time?

1.5 [3 points]
How much better is your new design for a client who will run BM5 for 50% of the time, BM1 for 20% of the time, and each of the other benchmarks for 10% of the time?

Problem 4: [15 points] Amdahl's Law

For this problem the dynamic instruction mix of a set of benchmark codes has been taken and there are 2 innovations that are under consideration for the new release of the processor. Again we will simply assume that all instructions take the same amount of time. Also assume that the clock frequency will not change due to either innovation. The following table shows the % of dynamic instructions in 4 classes: floating point, integer, memory, and other. One innovation primarily improves integer performance and the other targets floating point performance. Improvements in both memory and other with either innovation are more modest.

Instruction Mix	FP	Int	Mem	Other
	31%	41%	5%	23%
Baseline IPC	2.1	3.6	0.5	2.3
Innov1 IPC	3.5	3.6	0.7	2.6
Innov2 IPC	2.1	4.8	0.6	2.5

4.1 [5 points]
Which innovation is best and how much?

4.2 [5 points]
Assume that you can sell the new best innovation processor for an amount providing the equivalent performance per dollar as the baseline processor. **How much more expensive will the new processor be?**

4.3 [5 points]
Assume that the baseline processor costs $200 to build and that you can sell it for $250. If you were to combine the 2 innovations to get the best IPC numbers of either, but this would cause the clock speed to go down by 7% and the cost to manufacture to go up by $80 due to more expensive packaging requirements. **Based on performance per dollar, is the new combined innovation idea a good idea or a bad idea? Be sure to justify your answer.**