4. Amdahl's Law [40 points]

Consider the following three processors (X, Y, and Z) that are all of varying areas. Assume that the single-thread performance of a core increases with the square root of its area.

Processor X	**Processor Y**	**Processor Z**
Core Area = A	Core Area = $4A$	Core Area = $16A$

(a) You are given a workload where S fraction of its work is serial and $1 - S$ fraction of its work is **infinitely** parallelizable. If executed on a die composed of 16 Processor X's, what value of S would give a speedup of 4 over the performance of the workload on just Processor X?

$\frac{1}{(S+\frac{1-S}{16})} = 4$

$S = .2$

(b) Given a homogeneous die of area $16A$, which of the three processors would you use on your die to achieve maximal speedup? What is that speedup over just a single Processor X? Assume the same workload as in part(a).

Processor Y. Speedup of 5.
We know that using Processor X gives us a speedup of 4.
Since Processor Z would take up the full area and the performance increases with the square root of its area, we know that Processor Z gives us a speed up of 4.

Processor Y has a single-threaded speedup of 2 over Processor X.
$speedup = \frac{1}{(.2+\frac{.8}{4})} \times 2$
$speedup = 5$

We would use processor Y and get a speedup of 5.

(c) Now you are given a heterogeneous processor of area $16A$ to run the above workload. The die consists of 1 Processor Y and 12 Processor X's. When running the workload, all sequential parts of the program will be run on the larger core while all parallel parts of the program run exclusively on the smaller cores. What is the overall speedup achieved over a single Processor X?

Let n be the speedup of Processor Y over Processor X
$speedup = \frac{1}{\left(\frac{.2}{n}+\frac{.8}{16-n^2}\right)}$
$speedup = \frac{1}{\left(\frac{.2}{2}+\frac{.8}{12}\right)}$
$speedup = 6$

(d) One of the programmers decides to optimize the given workload so that it has 4% of its work in serial sections and 96% of its work in parallel sections. Which configuration would you use to run the workload if given the choices between the processors from part (a), part (b), and part (c)?

part (a) $speedup = \frac{1}{(.04+\frac{.96}{16})} = 10$
part (b) $speedup = \frac{1}{(.04+\frac{.96}{4})} \times 2 = 7.14$
part (c) $speedup = \frac{1}{(\frac{.04}{2}+\frac{.96}{12})} = 10$
You would use the processor from part (a) because it gives you the maximum speed up while being less complex to implement than the heterogeneous set up.

(e) What value of S would warrant the use of Processor Z over the configuration in part (c)?

Again, the speedup of Processor Z over one Processor X is 4. $\frac{1}{(\frac{S}{2}+\frac{1-S}{12})} = 4$
$S \geq .4$

(f) Typically, for a realistic workload, the parallel fraction is not infinitely parallelizable. What are the three fundamental reasons why?

1. Synchronization
2. Load imbalance
3. Resource contention

(g) Name a technique we discussed in class that takes advantage of the heterogeneous architecture to minimize the effect of one of the above reasons?

Accelerated Critical Sections or Bottleneck Identification and Scheduling