Chapter 1 Solutions

Case Study 1: Chip Fabrication Cost

1.1 a. Yield $= 1/(1 + (0.04 \times 2))^{14} = 0.34$

b. It is fabricated in a larger technology, which is an older plant. As plants age, their process gets tuned, and the defect rate decreases.

1.2 a. Phoenix:

$$\text{Dies per wafer} = \left(\pi \times (45/2)^2\right)/2 - (\pi \times 45)/\text{sqrt}(2 \times 2) = 795 - 70.7 = 724.5 = 724$$

$$\text{Yield} = 1/(1 + (0.04 \times 2))^{14} = 0.340$$

$$\text{Profit} = 724 \times 0.34 \times 30 = \$7384.80$$

b. Red Dragon:

$$\text{Dies per wafer} = \left(\pi \times (45/2)^2\right)/2 - (\pi \times 45)/\text{sqrt}(2 \times 1.2) = 1325 - 91.25 = 1234$$

$$\text{Yield} = 1/(1 + (0.04 \times 1.2))^{14} = 0.519$$

$$\text{Profit} = 1234 \times 0.519 \times 15 = \$9601.71$$

c. Phoenix chips: 25,000/724 = 34.5 wafers needed
Red Dragon chips: 50,000/1234 = 40.5 wafers needed

Therefore, the most lucrative split is 40 Red Dragon wafers, 30 Phoenix wafers.

1.3 a. Defect-free single core = Yield $= 1/(1 + (0.04 \times 0.25))^{14} = 0.87$
Equation for the probability that N are defect free on a chip:
#combinations $\times (0.87)^N \times (1 - 0.87)^{8-N}$

# defect-free	# combinations	Probability
8	1	0.32821167
7	8	0.39234499
6	28	0.20519192
5	56	0.06132172
4	70	0.01145377
3	56	0.00136919
2	28	0.0001023
1	8	4.3673E-06
0	1	8.1573E-08

Yield for Phoenix^4: $(0.39 + 0.21 + 0.06 + 0.01) = 0.57$
Yield for Phoenix^2: $(0.001 + 0.0001) = 0.0011$
Yield for Phoenix^1: 0.000004

b. It would be worthwhile to sell Phoenix^4. However, the other two have such a low probability of occurring that it is not worth selling them.

c.

$$\$20 = \frac{\text{Wafer size}}{\text{odd dpw} \times 0.28}$$

Step 1: Determine how many Phoenix4 chips are produced for every Phoenix8 chip.

There are 57/33 Phoenix4 chips for every Phoenix8 chip = 1.73

$$\$30 + 1.73 \times \$25 = \$73.25$$

Case Study 2: Power Consumption in Computer Systems

1.4 a. Energy: 1/8. Power: Unchanged.

b. Energy: $\text{Energy}_{new}/\text{Energy}_{old} = (\text{Voltage} \times 1/8)^2/\text{Voltage}^2 = 0.156$
Power: $\text{Power}_{new}/\text{Power}_{old} = 0.156 \times (\text{Frequency} \times 1/8)/\text{Frequency} = 0.00195$

c. Energy: $\text{Energy}_{new}/\text{Energy}_{old} = (\text{Voltage} \times 0.5)^2/\text{Voltage}^2 = 0.25$
Power: $\text{Power}_{new}/\text{Power}_{old} = 0.25 \times (\text{Frequency} \times 1/8)/\text{Frequency} = 0.0313$

d. 1 core = 25% of the original power, running for 25% of the time.

$$0.25 \times 0.25 + (0.25 \times 0.2) \times 0.75 = 0.0625 + 0.0375 = 0.1$$

1.5 a. Amdahl's law: 1/(0.8/4 + 0.2) = 1/(0.2 + 0.2) = 1/0.4 = 2.5

b. 4 cores, each at 1/(2.5) the frequency and voltage
Energy: $\text{Energy}_{quad}/\text{Energy}_{single} = 4 \times (\text{Voltage} \times 1/(2.5))^2/\text{Voltage}^2 = 0.64$
Power: $\text{Power}_{new}/\text{Power}_{old} = 0.64 \times (\text{Frequency} \times 1/(2.5))/\text{Frequency} = 0.256$

c. 2 cores + 2 ASICs vs. 4 cores

$$(2 + (0.2 \times 2))/4 = (2.4)/4 = 0.6$$

1.6 a. Workload A speedup: 225,000/13,461 = 16.7
Workload B speedup: 280,000/36,465 = 7.7
1/(0.7/16.7 + 0.3/7.7)

b. General-purpose: 0.70 × 0.42 + 0.30 = 0.594
GPU: 0.70 × 0.37 + 0.30 = 0.559
TPU: 0.70 × 0.80 + 0.30 = 0.886

c. General-purpose: 159 W + (455 W − 159 W) × 0.594 = 335 W
GPU: 357 W + (991 W − 357 W) × 0.559 = 711 W
TPU: 290 W + (384 W − 290 W) × 0.86 = 371 W

d.

Speedup	**A**	**B**	**C**
GPU	2.46	2.76	1.25
TPU	41.0	21.2	0.167
% Time	0.4	0.1	0.5

GPU: $1/(0.4/2.46 + 0.1/2.76 + 0.5/1.25) = 1.67$
TPU: $1/(0.4/41 + 0.1/21.2 + 0.5/0.17) = 0.33$

e. General-purpose: $14{,}000/504 = 27.8 \geq 28$
GPU: $14{,}000/1838 = 7.62 \geq 8$
TPU: $14{,}000/861 = 16.3 \geq 17$

d. General-purpose: $2200/504 = 4.37 \geq 4$, $14{,}000/(4 \times 504) = 6.74 \geq 7$
GPU: $2200/1838 = 1.2 \geq 1$, $14{,}000/(1 \times 1838) = 7.62 \geq 8$
TPU: $2200/861 = 2.56 \geq 2$, $14{,}000/(2 \times 861) = 8.13 \geq 9$

Exercises

1.7 a. Somewhere between 1.4^{10} and 1.55^{10}, or $28.9 - 80x$

b. 6043 in 2003, 52% growth rate per year for 12 years is 60,500,000 (rounded)

c. 24,129 in 2010, 22% growth rate per year for 15 years is 1,920,000 (rounded)

d. Multiple cores on a chip rather than faster single-core performance

e. $2 = x^4$, $x = 1.032$, 3.2% growth

1.8 a. 50%

b. Energy: $\text{Energy}_{new}/\text{Energy}_{old} = (\text{Voltage} \times 1/2)^2/\text{Voltage}^2 = 0.25$

1.9 a. 60%

b. $0.4 + 0.6 \times 0.2 = 0.58$, which reduces the energy to 58% of the original energy

c. newPower/oldPower = ½Capacitance × (Voltage × 0.8)^2 × (Frequency × 0.6)/½ Capacitance × Voltage × Frequency = $0.8^2 \times 0.6 = 0.256$ of the original power.

d. $0.4 + 0.3 \times 2 = 0.46$, which reduces the energy to 46% of the original energy

1.10 a. $10^9/100 = 10^7$

b. $10^7/10^7 + 24 = 1$

c. [need solution]

1.11 a. $35/10{,}000 \times 3333 = 11.67$ days

b. There are several correct answers. One would be that, with the current system, one computer fails approximately every 5 min. 5 min is unlikely to be enough time to isolate the computer, swap it out, and get the computer back on line again. 10 min, however, is much more likely. In any case, it would greatly extend the amount of time before 1/3 of the computers have failed at once. Because the cost of downtime is so huge, being able to extend this is very valuable.

c. $\$90{,}000 = (x + x + x + 2x)/4$
$\$360{,}000 = 5x$
$\$72{,}000 = x$
4th quarter = \$144,000/h

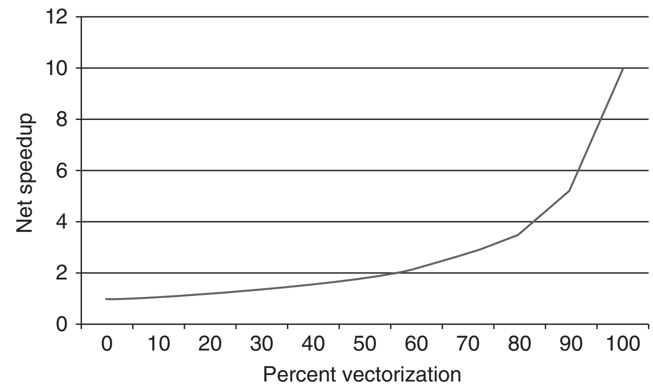

Figure S.1 Plot of the equation: $y = 100/((100 - x) + x/10)$.

1.12 a. See Figure S.1.

b. $2 = 1/((1 - x) + x/20)$
$10/19 = x = 52.6\%$

c. $(0.526/20)/(0.474 + 0.526/20) = 5.3\%$

d. Extra speedup with 2 units: $1/(0.1 + 0.9/2) = 1.82$. $1.82 \times 20 - 36.4$. Total speedup: 1.95. Extra speedup with 4 units: $1/(0.1 + 0.9/4) = 3.08$. $3.08 \times 20 - 61.5$. Total speedup: 1.97

1.13 a. old execution time $= 0.5$ new $+ 0.5 \times 10$ new $= 5.5$ new

b. In the original code, the unenhanced part is equal in time to the enhanced part (sped up by 10), therefore:
$(1 - x) = x/10$
$10 - 10x = x$
$10 = 11x$
$10/11 = x = 0.91$

1.14 a. $1/(0.8 + 0.20/2) = 1.11$

b. $1/(0.7 + 0.20/2 + 0.10 \times 3/2) = 1.05$

c. fp ops: $0.1/0.95 = 10.5\%$, cache: $0.15/0.95 = 15.8\%$

1.15 a. $1/(0.5 + 0.5/22) = 1.91$

b. $1/(0.1 + 0.90/22) = 7.10$

c. $41\% \times 22 = 9$. A runs on 9 cores. Speedup of A on 9 cores: $1/(0.5 + 0.5/9) = 1.8$ Overall speedup if 9 cores have 1.8 speedup, others none: $1/(0.6 + 0.4/1.8) = 1.22$

d. Calculate values for all processors like in c. Obtain: 1.8, 3, 1.82, 2.5, respectively.

e. $1/(0.41/1.8 + 0.27/3 + 0.18/1.82 + 0.14/2.5) = 2.12$

1.16 a. $1/(0.2 + 0.8/N)$

b. $1/(0.2 + 8 \times 0.005 + 0.8/8) = 2.94$

c. $1/(0.2 + 3 \times 0.005 + 0.8/8) = 3.17$

d. $1/(.2 + \log N \times 0.005 + 0.8/N)$

e. $d/dN\ (1/((1 - P) + \log N \times 0.005 + P/N) = 0)$