Karan Vora (N12229954) (kv2154)
Computer Systems Architecture Assignment 2

==

Problem 1):

$\text{Clockrate}_{\text{Optimised}}$ = **x**
$\text{Clockrate}_{\text{Unoptimised}}$ = 1.05 * **x**

Load/Store IC of Unoptimized = 0.3 * IC
Dynamic IC of Unoptimized = 0.7 * IC

Load/Store IC of Optimized = 2/3 * Load/Store IC of Unoptimized = 0.2 * IC
Dynamic IC of Optimized = Dynamic IC of Unoptimized = 0.7 * IC

Now, From the formula of CPU Time

CPUTime = $\Sigma(IC_i)$ * 1/Clockrate

$$CPUTime_{Unoptimized} = (0.3*IC+0.7*IC)*(\frac{1}{1.05*x})$$

$$CPUTime_{Optimized} = (0.2*IC+0.7*IC)*(\frac{1}{x})$$

$$Speed\ up = \frac{CPUTime_{Optimized}}{CPUTime_{Unoptimized}}$$

$$\frac{CPUTime_{Optimized}}{CPUTime_{Unoptimized}} = \frac{0.9*IC*(\frac{1}{x})}{IC*(\frac{1}{1.05*x})}$$

$$CPUTime_{Optimized} = 0.945*CPUTime_{Unoptimized}$$

From above mention relationship, we can conclude that the optimized CPU is faster than unoptimized version.

==

Problem 2):

Solution A):

Throughput of GPU for A = 13,461
Throughput of TPU for A = 225,000

Speedup of TPU over GPU in A = $\frac{Throughput\ of\ TPU\ for\ A}{Throughput\ of\ GPU\ for\ A} = \frac{225{,}000}{13{,}461} = 16.715$

Throughput of GPU for B = 36,465
Throughput of TPU for B = 280,000

Speedup of TPU over GPU in B = $\frac{Throughput\ of\ TPU\ for\ A}{Throughput\ of\ GPU\ for\ B} = \frac{280{,}000}{36{,}465} = 7.68$

Time spends for A = 0.7 * **x**
Time spends for B = 0.3 * **x**

ET of TPU in A * 16.715 = ET of GPU in A = 0.7 * T
ET of TPU in B * 7.68 = ET of GPU in B = 0.3 * T

Overall Speedup = $\frac{Total\ ET\ of\ GPU}{Total\ ET\ of\ TPU}$

$$Overall\ Speedup = \frac{T}{\left(\frac{0.7}{16.715} + \frac{0.3}{7.68}\right) * T}$$

Overall Speedup of TPU over GPU = 12.3547

Solution B):

General purpose IPS for A = 42%
General purpose IPS for B = 100%
Max General purpose IPS = 0.42*0.7 + 1*0.3 = 0.594 = 59.4%

GPU IPS for A = 37%
GPU IPS for B = 100%
Max GPU IPS = 0.37*0.7 + 0.3*1 = 0.559 = 55.9%

TPU IPS for A = 80%
TPU IPS for B = 100%
Max TPU IPS = 0.8*0.7 = 0.3*1 = 0.86 = 86%

Solution C):

Performance per watt = $\frac{Throughput}{Watt}$

Performance for watt of TPU over GPU = $\frac{\left(\frac{Throughput\ of\ TPU}{Watt\ of\ TPU}\right)}{\left(\frac{Throughput\ of\ GPU}{Watt\ of\ GPU}\right)}$

$= \frac{\left(\frac{0.86}{384-290}\right)}{\left(\frac{0.559}{991-357}\right)} = \frac{545.24}{52.25} = 10.367$

Solution D):

General purpose Throughput of A = 5482
General purpose Utilization of A = 0.4 * 5482 = 2192.8

General purpose Throughput of B = 13194
General purpose Utilization of B = 0.1 * 13194 = 1319.4

General purpose Throughput of C = 12000
General purpose Utilization of C = 0.5 * 12000 = 6000

GPU Throughput of A = 13461
GPU Utilization of A = 0.4 * 13194 = 5277.6

GPU Throughput of B = 36465
GPU Utilization of B = 0.1 * 36465 = 3646.5

GPU Throughput of C = 15000
GPU Utlization of C = 0.5 * 15000 = 7500

TPU Throughput of A = 225000
TPU Utilization of A = 0.4 * 225000 = 90000

TPU Throughput of B = 280000
TPU Utilization of B = 0.1 * 280000 = 28000

TPU Throughput of C = 2000
TPU Utilization of C = 1000

Speed up of GPU over General Purpose

GPU Time = $\frac{2192.8}{5277.6}+\frac{1319.4}{3646.5}+\frac{6000}{7500}=1.57732$

GPU Speedup = $\frac{1}{1.57732}=0.633987$

Speedup of TPU over General Purpose

TPU Time = $\frac{2192.8}{90000}+\frac{1319.4}{28000}+\frac{6000}{1000}=6.07149$

TPU Speedup = $\frac{1}{6.07149}=0.164704$

Solution E):

Total Capacity = 14000 Watt

TDP of CPU = 504,
Number of CPU can be cooled = 14000/504 = 27.78 ≈ 27

TDP of GPU = 1838,
Number of GPU can be cooled = 14000/1838 = 7.6 ≈ 7

TDP of GPU = 861
Number of TPU can be cooled = 14000/861 = 16.26 ≈ 16

Solution F):

Cooling capacity per sqft = 200 Watt
Total area = 11 sqft
Total cooling capacity = 2200 Watt

TDP of CPU = 504
Number of CPU can be cooled = 2200/504 = 4.36 ≈ 4

TDP of GPU = 1838
Number of GPU can be cooled = 2200/1838 = 1.19 ≈ 1

TDP of TPU = 861
Number of TPU can be cooled = 2200/861 = 2.55 ≈ 2

Now 1 Cooling door can cool upto,
General Purpose CPU = 27
GPU = 7

TPU = 16
So realistically, 11 sqft area can easily cool the components mentioned using 1 cooling door.

==

Problem 3):

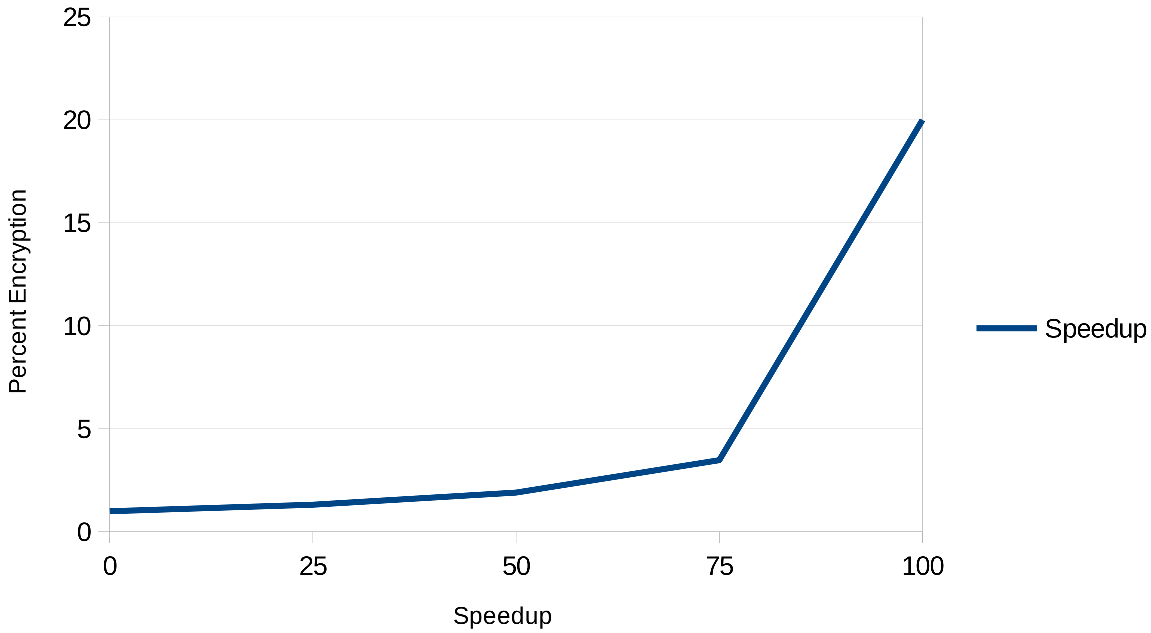
Solution A):

Let the percent of encryption be X and speedup enhance is given as 20
Formula for Speedup is

$$Speedup=\frac{1}{\left(1-\frac{X}{100}\right)+\left(\frac{X}{20*100}\right)}$$

$$Speedup=\frac{2000}{2000-19*X}$$

Percent Encryption	Speedup
0	1
25	1.31148
50	1.90476
75	3.47826
100	20



Solution B):

Speedup = 2

$$2=\frac{2000}{2000-19*X}$$

X = 52.6316 %

Solution C):
Percent encryption when speedup is 2 = 52.6316
Percent non-encryption when speedup is 2 = 47.3684

so if the original time is 100 seconds without speed, encryption takes 52.6316 S and non encryption takes 47.3684 S

When theres speed up of 2, the time will be 50, so encryption will be 50 – 47.3684 = 2.63158
so with speed up of 2, S are spent in encryption mode which is 5.26316

Percent spent in encryption mode = 5.26316%

==

Problem 4):

Solution A):

$$\frac{ExecutionTime_{Old}}{ExecutionTime_{New}} = \frac{1}{(1-F)+(\frac{F}{Speedup})}$$

$$= \quad (1-0.5)+(\frac{0.5}{0.1}) = 0.5+5 = 5.5$$

Overall Speedup = 5.5

Solution B):

$$\frac{ExecutionTime_{New}}{ExecutionTime_{old}} = (1-F_{Enhanced})+(\frac{F_{Enhanced}}{SpeedUp_{Enhanced}})$$

$$= \quad (1-F_{Enhanced})+(\frac{F_{Enhanced}}{10}) = 1-F_{Enhanced}+(0.9*F_{Enhanced}) = 1-(0.9*F_{Enhanced})$$

Now,

$$1-(0.9*F_{Enhanced}) = \frac{1}{5.5}$$

Thefore,

$$F_{Enhanced} = 0.91$$

Percent of time in new mode = 91%

==

Problem 5):

$$Speedup_{Overall} = \frac{ExecutionTime_{Old}}{ExecutionTime_{New}} = \frac{1}{(1-Fraction_{Enhanced})+(\frac{Francrion_{Enhanced}}{Speedup_{Enhanced}})}$$

Solution A):

Parallelizable = 80%
Let the processors be N

$$Speedup_{Overall} = \frac{1}{0.2+\frac{0.8}{N}}$$

Solution B):

Number of processors = 8
Communication overhead = 0.5% of Execution time

$$Speedup_{Overall} = \frac{1}{0.2 + (8 * 0.005) + \frac{0.8}{8}} = 2.94$$

Solution C):

Number of processors = 8
Everytime the processor count is doubled, communication overhead is increased by 0.5% of total execution time.
Times the processor count doubled = 3 (1 → 2, 2 → 4, 4 → 8)

Therefore,

$$Speedup_{Overall} = \frac{1}{0.2 + (3 * 0.005) + \frac{0.8}{8}} = 3.17$$

Solution D):

Number of processors = N
Everytime the processor count is doubled, communication overhead is increased by 0.5% of total execution time.

Now,

N is number of processor we get after doubling the count

$$2^n = N$$

$$n \log 2 = \log N$$

$$n = \frac{\log N}{\log 2}$$

Therefore,

$$Speedup_{Overall} = \frac{1}{0.2 + \left(\frac{\log N}{\log 2} * 0.005\right) + \frac{0.8}{N}}$$

Solution E):

Let the percent of parallelization = P
Number of processors = N
Everytime the processor count is doubled, communication overhead is increased by 0.5% of total execution time.
Number of steps to double the count $n = \frac{logN}{\log 2}$

Therefore, General equation of speedup

$$Speedup_{Overall} = \frac{1}{(1-P) + (\frac{logN}{\log 2} * 0.005) + (\frac{P}{N})}$$

==

Problem 6):

Solution A):

Percent of parallelization of A = 50%
Therefore speedup when running A in parallel vs serial = 50

Solution B):

Percent of parallelization of D = 90%
Therefore speedup when running D in parallel vs serial = 90

Solution C):

Cores used by A = 41% of 22 = 9 Cores

$$Speedup\ of\ A = 0.41 * (\frac{0.5}{9} + 0.5) = 0.228$$

$$Speedup_{Overall} = \frac{1}{0.228 + 0.27 + 0.18 + 0.14} = \frac{1}{0.818} = 1.22$$

Solution D):

Cores used by A = 41% of 22 = 9 Cores

$$Speedup\ of\ A = 0.41 * (\frac{0.5}{9} + 0.5) = 0.228$$

Cores used by B = 27% of 22 = 6 Cores

$$Speedup\ of\ B = 0.27*\left(\frac{0.8}{6}+0.2\right) = 0.0899$$

Cores used by C = 18% of 22 = 4 Cores

$$Speedup\ of\ C = 0.18*\left(\frac{0.6}{4}+0.4\right) = 0.099$$

Cores used by D = 14% of 22 = 3 Cores

$$Speedup\ of\ D = 0.14*\left(\frac{0.9}{4}+0.1\right) = 0.056$$

Now,

$$Speedup_{Overall} = \frac{1}{0.228+0.0899+0.099+0.056} = 2.115$$

Solution E):

Cores used by A = 41% of 22 = 9 Cores

$$Speedup\ of\ A = 0.41*\left(\frac{0.5}{9}+0.5\right) = 0.228 = 22.8\%$$

Cores used by B = 27% of 22 = 6 Cores

$$Speedup\ of\ B = 0.27*\left(\frac{0.8}{6}+0.2\right) = 0.0899 = 8.99\%$$

Cores used by C = 18% of 22 = 4 Cores

$$Speedup\ of\ C = 0.18*\left(\frac{0.6}{4}+0.4\right) = 0.099 = 9.9\%$$

Cores used by D = 14% of 22 = 3 Cores

$$Speedup\ of\ D = 0.14*\left(\frac{0.9}{4}+0.1\right) = 0.056 = 5.6\%$$

===

Problem 7):

Solution A):

Floating point fraction = 20%
Speedup = 2

$$Speedup_{Overall} = \frac{1}{(1 - Fraction) + \frac{Fraction}{Speedup}} = \frac{1}{(1 - 0.20) + \frac{0.20}{2}} = 1.11$$

Solution B):

Speedup = 2/3
Data cache = 10% of execution time

$$Speedup_{Overall} = \frac{1}{0.7 + \frac{0.2}{2} + (\frac{0.1 * 3}{2})} = 1.05$$

Solution C):

Percentage time spent in Floating point = $\frac{(\frac{0.2}{2})}{0.7 + \frac{0.2}{2} + (0.1 * 1.5)} = 0.105 = 10.5\,\%$

Percentage time spent in Data cache access = $\frac{(0.1 * 1.5)}{0.7 + \frac{0.2}{2} + (0.1 * 1.5)} = 0.1579 = 15.79\,\%$

===

Problem 8):

Solution A):

Floating point fraction = 20%
Speedup = 2

$$Speedup_{Overall} = \frac{1}{(1 - Fraction) + \frac{Fraction}{Speedup}} = \frac{1}{(1 - 0.20) + \frac{0.20}{2}} = 1.11$$

Solution B):

Speedup = 2/3
Data cache = 10% of execution time

$$Speedup_{Overall}=\frac{1}{0.7+\frac{0.2}{2}+(\frac{0.1*3}{2})}=1.05$$

Solution C):

Percentage time spent in Floating point = $$\frac{(\frac{0.2}{2})}{0.7+\frac{0.2}{2}+(0.1*1.5)}=0.105=10.5\%$$

Percentage time spent in Data cache access = $$\frac{(0.1*1.5)}{0.7+\frac{0.2}{2}+(0.1*1.5)}=0.1579=15.79\%$$