Shahmir Shehzad

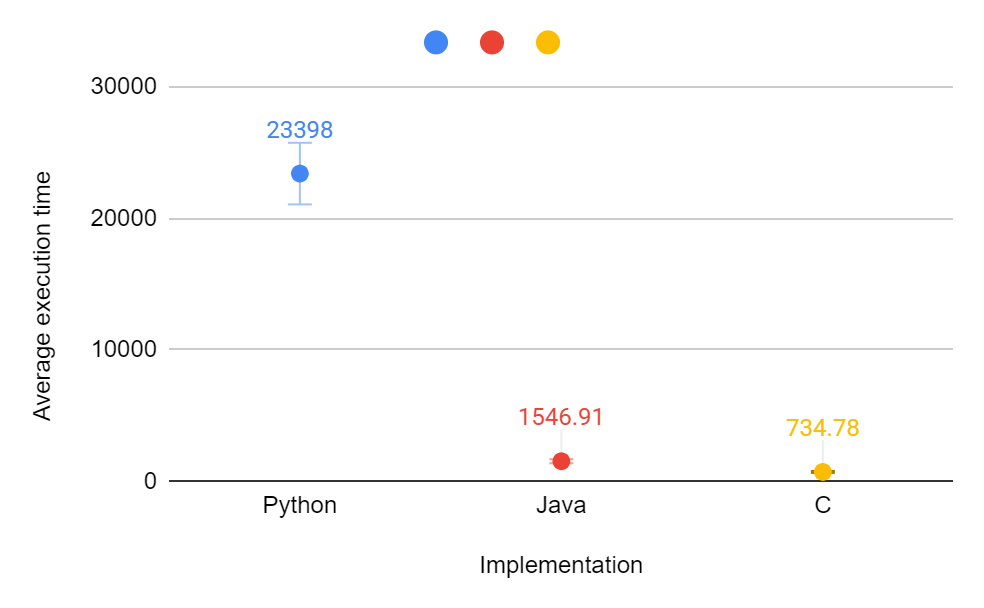
PDC - 6C

Assignment-2

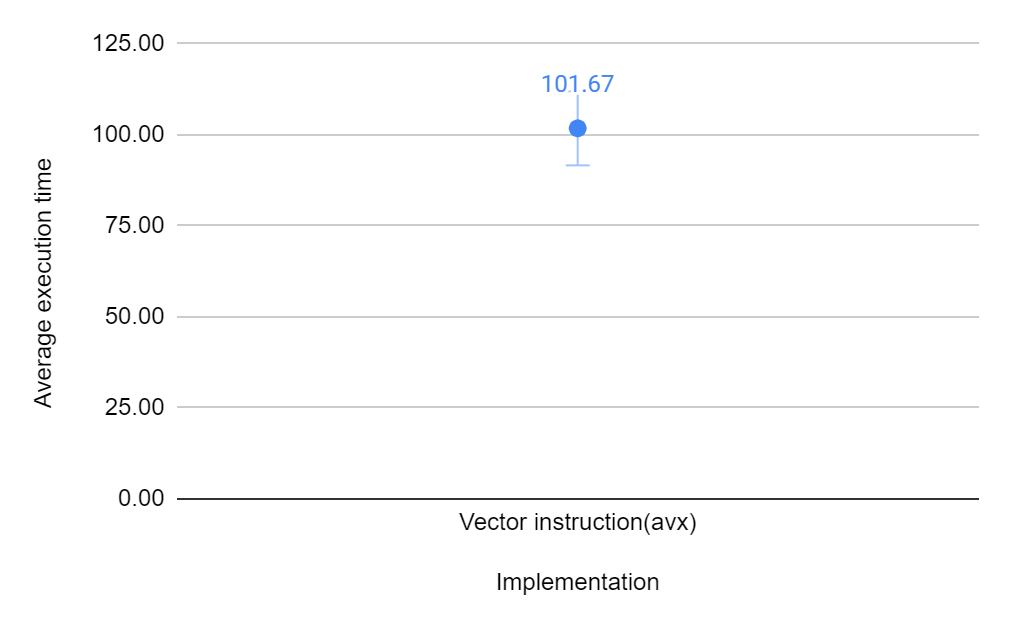
**Graphs and tables from assignment 1**

|  |  |  |
| --- | --- | --- |
| Implementation | Average running time(s) | Speedup |
| Python | 23,398.19 | 1x |
| Java | 1546.9 | 15x |
| C | 734.77 | 32x |
| Pthreads in C | 319.7 | 73x |
| AVX in C | 102 | 229x |
| Pthreads + AVX in C | 62 | 377x |

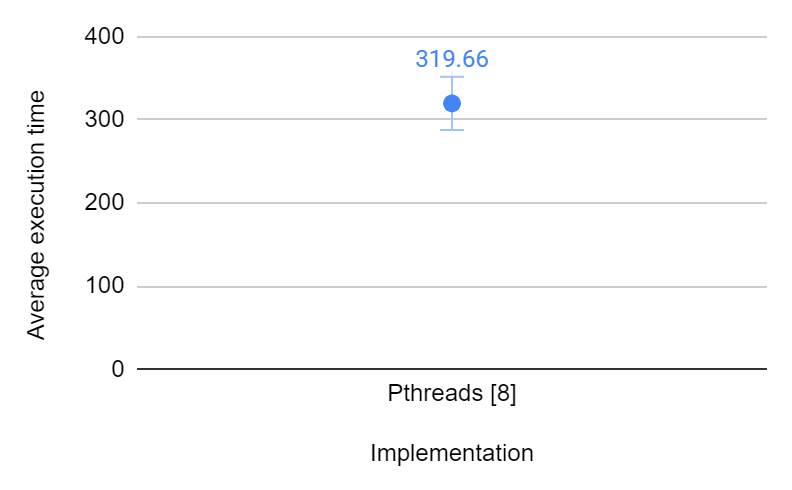
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Implementation | Run 1 | Run 2 | Run 3 | Avg | SD |
| Python | 21698.53 | 24497.84 | 23997.63 | 23398 | 1492.88 |
| Java | 1496.57 | 1535.75 | 1608.4 | 1546.91 | 56.74 |
| C | 733.94 | 731.7 | 738.7 | 734.78 | 3.57 |



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Implementation | Run 1 | Run 2 | Run 3 | Avg | SD |
| Vector instruction(avx) | 129 | 101 | 75 | 101.67 | 27.00617213 |

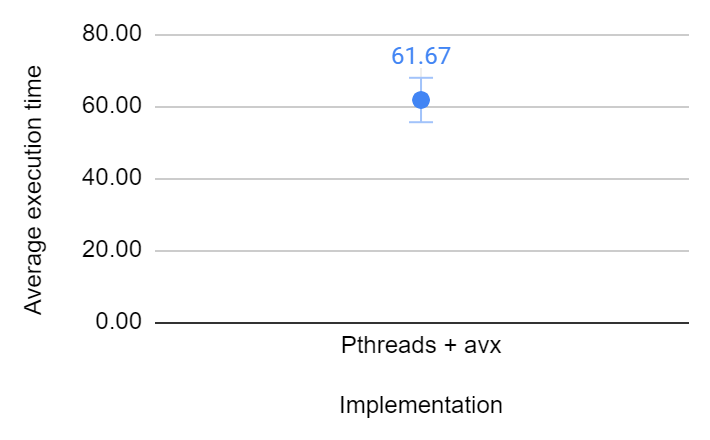
****

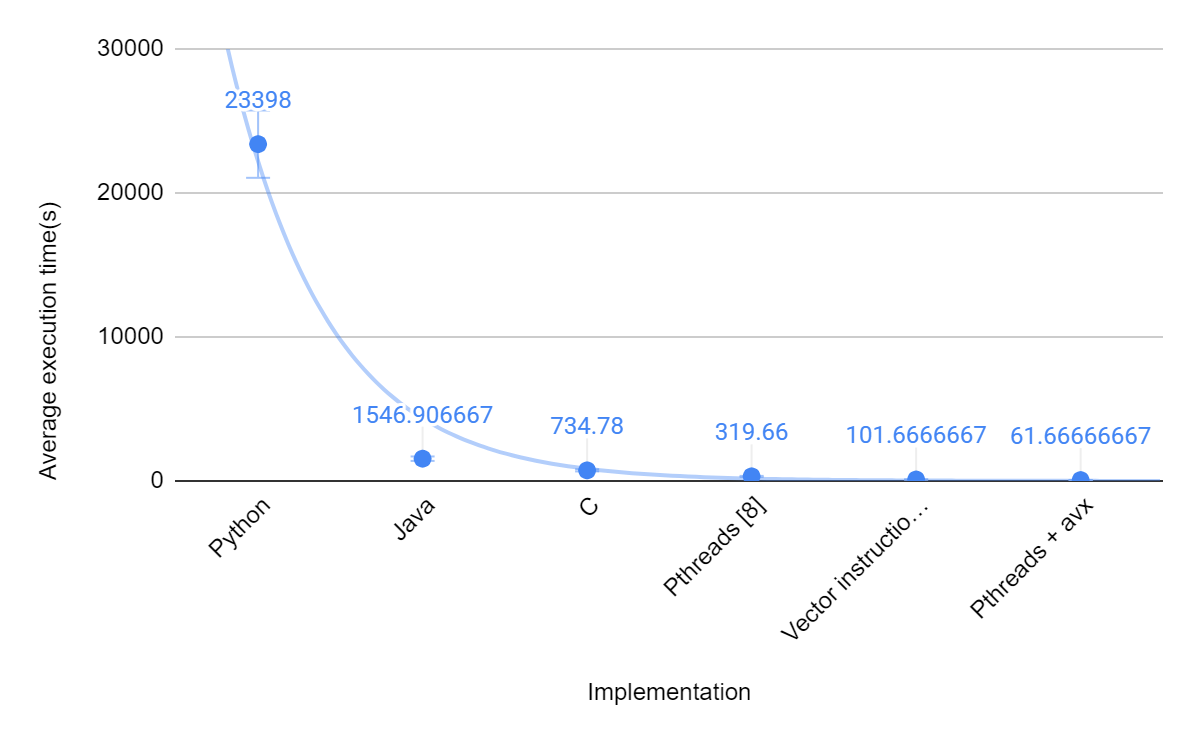
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Implementation | Run 1 | Run 2 | Run 3 | Avg | SD |
| Pthreads [8] | 350.61 | 331.76 | 276.61 | 319.66 | 38.46 |

****

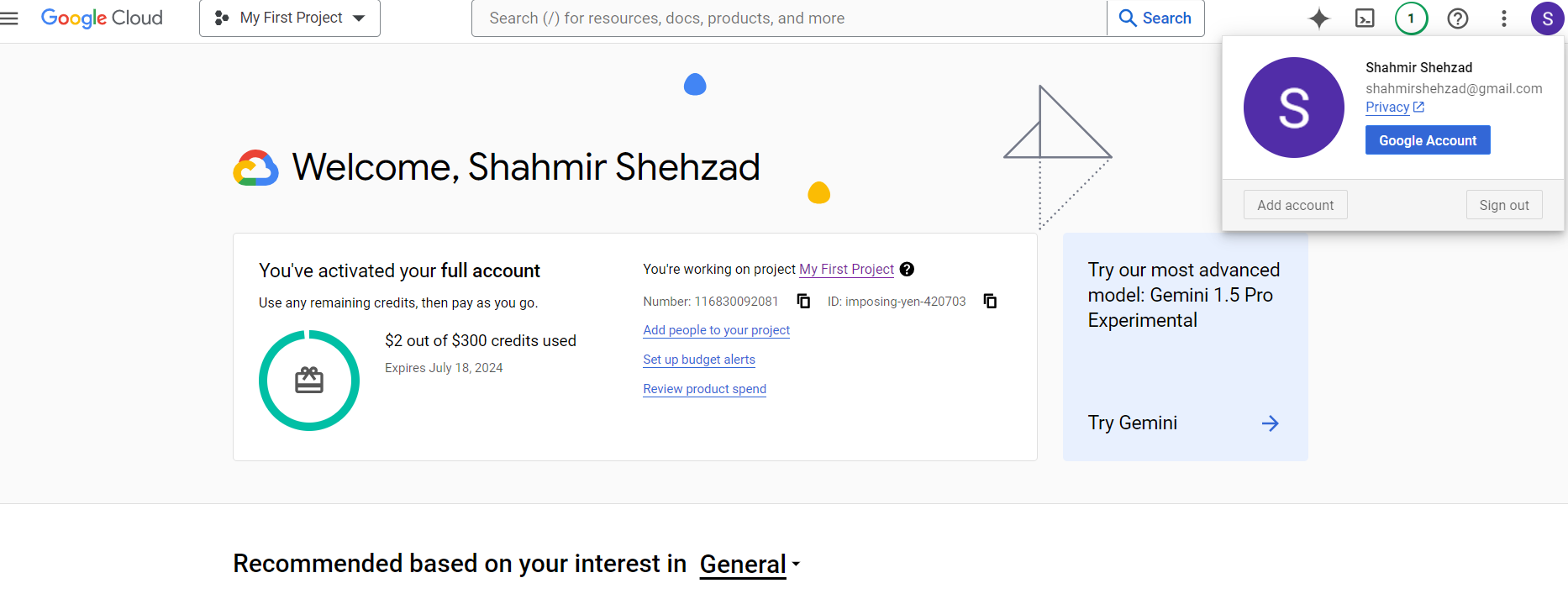
**Part 4**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Implementation | Run 1 | Run 2 | Run 3 | Avg | SD |
| Pthreads + avx | 74 | 47 | 64 | 61.67 | 13.65039682 |

****

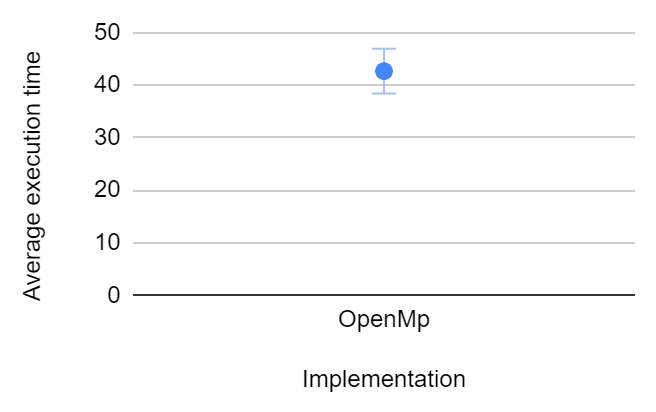
****

**Part 0**

****

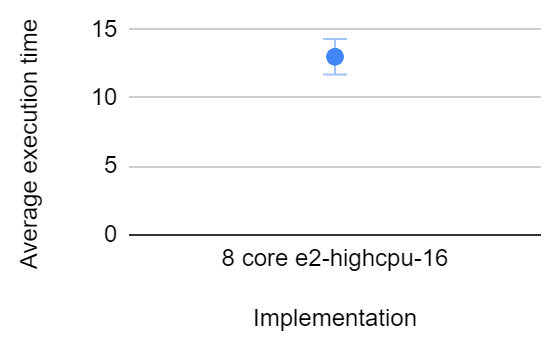
**Part 1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Implementation | Run 1 | Run 2 | Run 3 | Avg | SD |
| OpenMp | 36 | 44 | 48 | 42.67 | 6.11 |

****

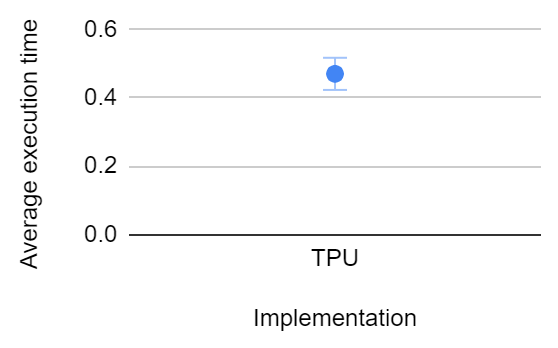
**Part 2**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Implementation | Run 1 | Run 2 | Run 3 | Avg | SD |
| 16 core e2-standard | 13 | 13 | 13 | 13 | 0.00 |

****

**Part 3**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Implementation | Run 1 | Run 2 | Run 3 | Avg | SD |
| TPU | 0.51 | 0.44 | 0.46 | 0.47 | 0.04 |

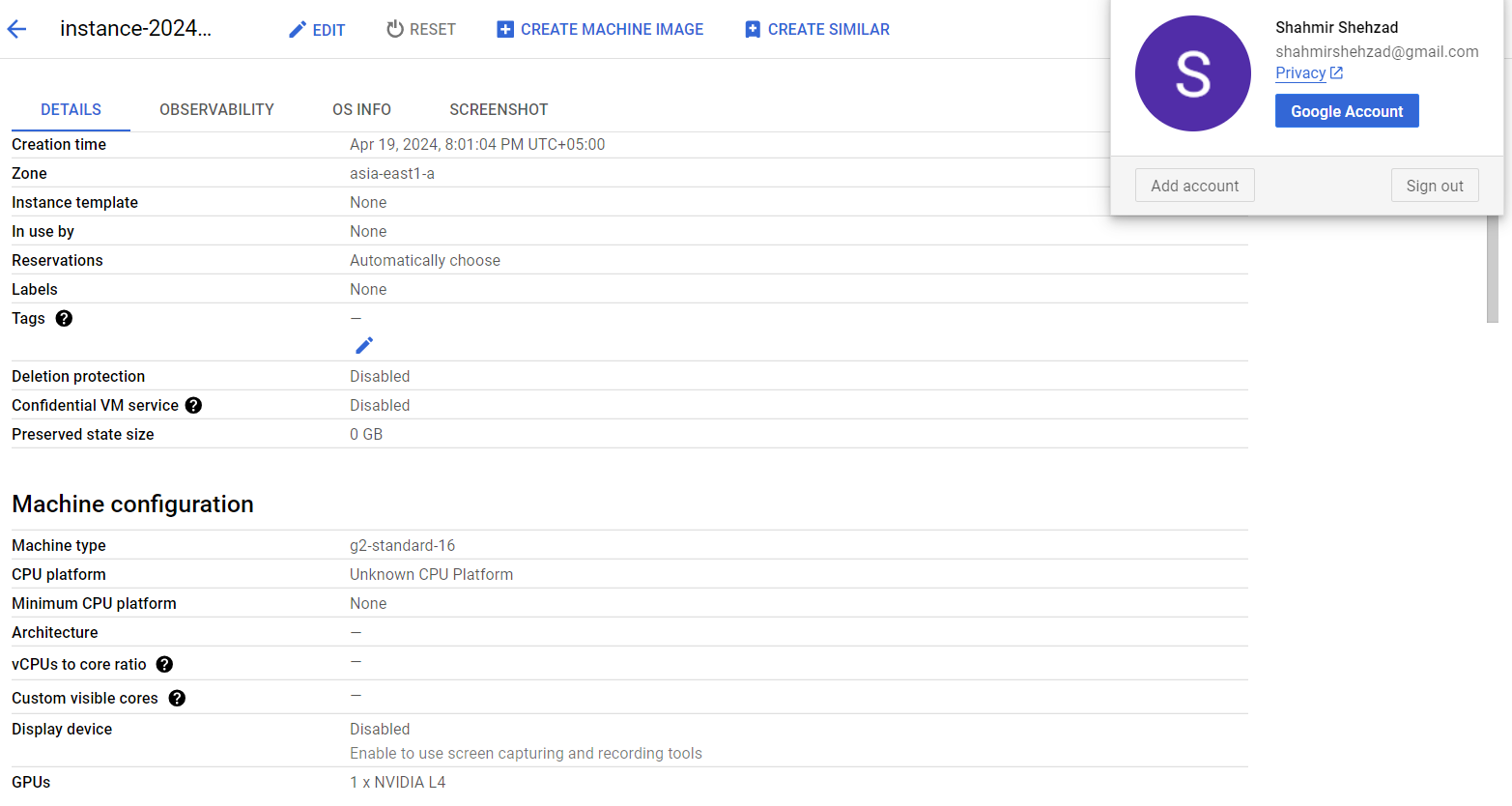
****

**Part 4**

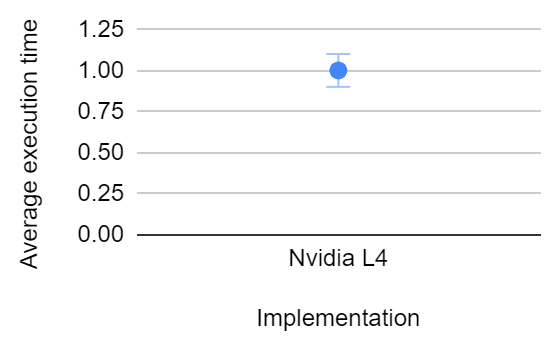
**Gpu : 1 x NVIDIA L4**

**Region : asia-east1-a**

**Proof :**

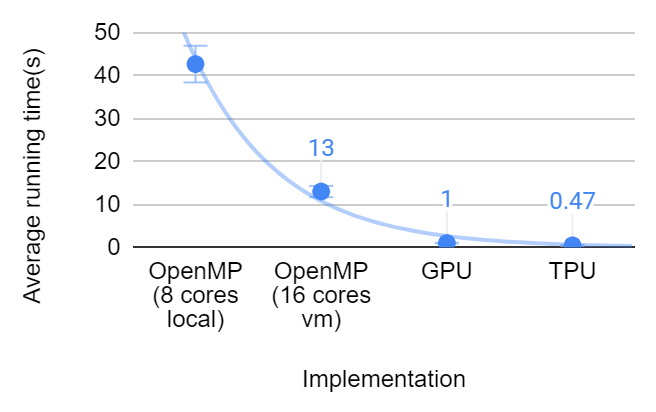
****

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Implementation | Run 1 | Run 2 | Run 3 | Avg | SD |
| Nvidia L4 | 1 | 1 | 1 | 1 | 0.00 |

****

**Part 5**

|  |  |  |
| --- | --- | --- |
| Implementation | Average running time(s) | Speedup |
| OpenMP (8 cores local) | 42.67 | 1x |
| OpenMP (16 cores vm) | 13 | 3x |
| GPU | 1 | 43x |
| TPU | 0.47 | 91x |

****

TPU gave the best results. TPU uses systolic arrays that consist of processing elements (PE). Each PE performs matrix multiplication and accumulates the result and passes it onto neighboring PE in the next pipeline iteration. This approach uses both parallelism and pipelining.

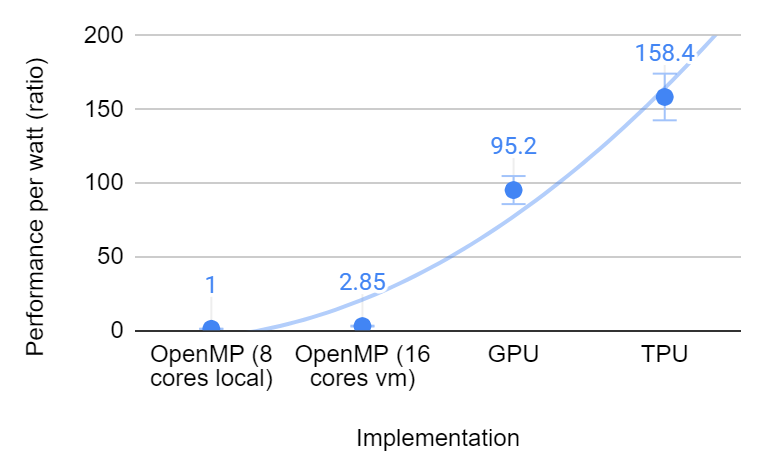
**Part 6**

**Performance per watt = No. of operations / ( Avg execution time \* Power )**

**Assuming n = 19.2 x 10^9 (Gflop) for cpu**

**Assuming n = 67 x 10^9 (Gflop) for gpu, tpu**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Implementation | Average execution time(s) | Power consumed(W) | Performance per watt | Performance per watt (ratio) |
| OpenMP (8 cores local) | 42.67 | 46 | 9781844 | 1 |
| OpenMP (16 cores vm) | 13 | 53 | 27866473.2 | 2.85 |
| GPU | 1 | 72 | 930555555.6 | 95.2 |
| TPU | 0.47 | 92 | 1549491212 | 158.4 |

****