

Edwin Tomy George

CS 1190 Parallel Computing

Instructor: David Pruitt

Wednesday, February 23rd, 2022

Assignment III Report: Map Reduce

Issues, Solutions & Time Management

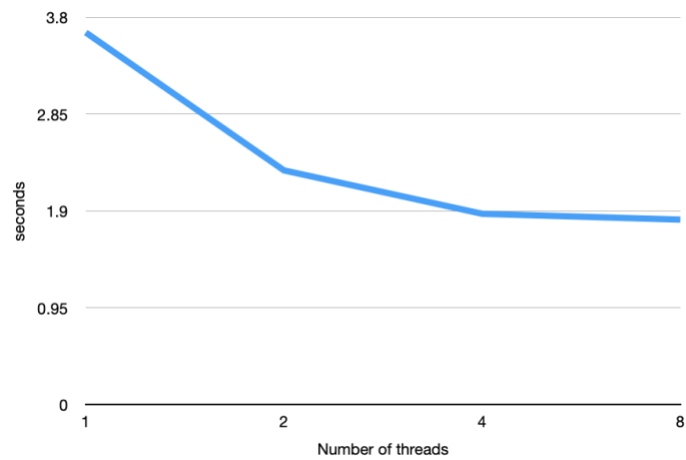
During the completion of the assignment, there were issues with the global dictionary which stored the counts of the words. This was solved by introducing local dictionary variables inside the parallel section. Race conditions also occurred when adding the local dictionaries back, which was resolved using locks during the addition.

Overall, it took me around 3 hours to complete the assignment, from which 1 hour was spend debugging and half an hour on the report.

Number of Threads-Based Performance

The performance increased modestly from one-thread to two-threads. Once we add more threads, the performance improvement becomes even more insignificant. This may be due because most of the capable improvement was already achieved with two threads, and adding more threads increases computational resources, including reading the files.

# Threads	word count
1	3.6521 s
2	2.2988 s
4	1.8730 s
<u>8</u>	1.8158 s



Conclusions

Overall, the assignment showcased the importance of creating good algorithms for faster performance, since it can have much greater improvements than parallelization. Better algorithms can also help improve parallel programming.

Output of cplInfoDump.sh

model name : Intel(R) Core(TM) i5-1038NG7 CPU @ 2.00GHz

2 18 11