Alex Adrian Avila

Parallel Computing

Project 3 Map Reduce MPI

**Introduction**

For this project our task was the same as the previous assignment to read many files and count the number appearances of certain words that they appear throughout the whole set of files. However, this time we will use the MPI library instead of pymp which uses processes that communicate through processes.

**Problems**

The assignment was a bit easier than I thought however I ran into problems that were basic. I got stock for a while because I was not able to get the exact number of words that we were supposed to get, I thought I was not splitting the list correctly but what happened is that I was using the indices of a list and getting the elements of another file list.

**Performance**

MPI results

Chart, line chart

Description automatically generated

PYMP results

Chart, line chart

Description automatically generated

**Analysis**

It looks like using processes is way slower than the other method even though is useful not having shared memory to avoid errors sending messages does not seem as effective as using threads and shared memory for this problem.

**Conclusion**

Using processes seems useful if you probably have more computers that would talk to each other and a lot of memory to store all the process memory allocation but it seems like if the problem is a bit smaller the thread could be a better solution.

**cpuDump**

model name : Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz

4 36 220