

Parallel Map Reduce

Our assignment was to create a parallel program using pypm that would count the total number of instances for a set of words in multiple documents. I encountered a couple errors mostly some weird issues when declaring the shared dictionary outside of the function and I had also had some problems using findall(). No bugs to report, the program is functioning. This assignment took me about 5 hours which included taking the time to learn more about regular expressions.

1 thread - average time was 0.39228

2 thread - average time was 0.34187

4 thread - average time was 0.27606

8 thread - average time was 0.23519

As you can see from the results the program ran faster as the number of threads increased. The speed almost doubled when comparing 1 thread to 8.

CPU Info

```
student@linux-4oae:~> dumpCPUInfo.sh cpuinfo.txt
model name      : Intel(R) Core(TM) i5-4278U CPU @ 2.60GHz
4              36              216
```