Sistemas concurrentes
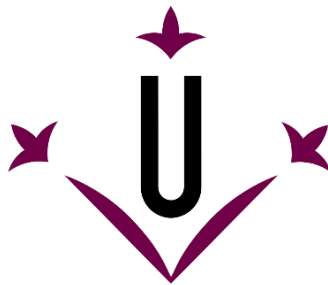
Concurrencia MapReduce en C++

Universitat de Lleida

José Almunia

Albert Giné

Introducción

Esta práctica tiene como objetivo la implementación de una aplicación concurrente multihilo utilizando pthreads con el lenguaje de programación C++.
Para ello haremos uso de los recursos aprendidos en clase.
El equipo donde hemos ejecutado el programa para recopilar los distintos tiempos de ejecución tiene el siguiente hardware:
CPU: AMD Ryzen 5 2600 (6 nucleos 12 hilos con OC a 4,1GHz)
RAM: Corsair Vengeance 2x8GB (OC a 3466MHz)

Ejecución secuencial:

time ./WordCount ./Test ./Output

```
Processing input file 41-0.txt
Processing input file all.txt
Processing input file test1.txt
Processing input file pg5200.txt
Processing input file Beowulf.txt
Processing input file Quijote.txt
Processing input file 74-0.txt
Processing input file pg42.txt
Processing input file pg20203.txt
Processing input file test12.txt
Processing input file 16-0.txt

real    0m17.456s
user    0m16.683s
sys     0m0.768s
```

Ejecuciones concurrentes:

time ./WordCount ./Test ./Output 1

```
Thread number 0, processing input file ./Test//all.txt
Thread number 1, processing input file ./Test//test1.txt
Thread number 2, processing input file ./Test//pg5200.txt
Thread number 3, processing input file ./Test//Beowulf.txt
Thread number 4, processing input file ./Test//Quijote.txt
Thread number 5, processing input file ./Test//74-0.txt
Thread number 6, processing input file ./Test//pg20203.txt
Thread number 8, processing input file ./Test//test12.txt
Thread number 9, processing input file ./Test//16-0.txt
Thread number 10, processing input file ./Test//16-0.txt
Thread number 7, processing input file ./Test//16-0.txt
Thread number 10, executing reduce phase

real    0m1.816s
user    0m3.058s
sys     0m0.558s
```

time ./WordCount ./Test ./Output 2

```
Thread number 0, processing input file ./Test//all.txt
Thread number 1, processing input file ./Test//test1.txt
Thread number 2, processing input file ./Test//pg5200.txt
Thread number 3, processing input file ./Test//Beowulf.txt
Thread number 4, processing input file ./Test//Quijote.txt
Thread number 5, processing input file ./Test//74-0.txt
Thread number 6, processing input file ./Test//pg42.txt
Thread number 7, processing input file ./Test//pg20203.txt
Thread number 8, processing input file ./Test//test12.txt
Thread number 9, processing input file ./Test//16-0.txt
Thread number 10, processing input file ./Test//16-0.txt
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase

real    0m1.388s
user    0m2.308s
sys     0m0.481s
```

time ./WordCount ./Test ./Output 3

```
Thread number 0, processing input file ./Test//all.txt
Thread number 1, processing input file ./Test//test1.txt
Thread number 2, processing input file ./Test//pg5200.txt
Thread number 3, processing input file ./Test//Beowulf.txt
Thread number 4, processing input file ./Test//Quijote.txt
Thread number 7, processing input file ./Test//test12.txt
Thread number 6, processing input file ./Test//test12.txt
Thread number 5, processing input file ./Test//16-0.txt
Thread number 9, processing input file ./Test//16-0.txt
Thread number 10, processing input file ./Test//16-0.txt
Thread number 8, processing input file ./Test//16-0.txt
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase

real    0m1.255s
user    0m1.995s
sys     0m0.472s
```

time ./WordCount ./Test ./Output 4

```
Thread number 3, processing input file ./Test//Beowulf.txt
Thread number 4, processing input file ./Test//74-0.txt
Thread number 9, processing input file ./Test//16-0.txt
Thread number 8, processing input file ./Test//16-0.txt
Thread number 7, processing input file ./Test//16-0.txt
Thread number 0, processing input file ./Test//16-0.txt
Thread number 5, processing input file ./Test//16-0.txt
Thread number 2, processing input file ./Test//16-0.txt
Thread number 1, processing input file ./Test//16-0.txt
Thread number 6, processing input file ./Test//16-0.txt
Thread number 10, processing input file ./Test//16-0.txt
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase

real    0m0.393s
user    0m2.034s
sys     0m0.480s
```

time ./WordCount ./Test ./Output 5

```
Thread number 0, processing input file ./Test//all.txt
Thread number 1, processing input file ./Test//test1.txt
Thread number 2, processing input file ./Test//pg5200.txt
Thread number 3, processing input file ./Test//Beowulf.txt
Thread number 5, processing input file ./Test//74-0.txt
Thread number 6, processing input file ./Test//pg42.txt
Thread number 7, processing input file ./Test//pg20203.txt
Thread number 8, processing input file ./Test//test12.txt
Thread number 4, processing input file ./Test//16-0.txt
Thread number 9, processing input file ./Test//16-0.txt
Thread number 10, processing input file ./Test//16-0.txt
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase
Thread number 10, executing reduce phase

real    0m14.318s
user    0m15.843s
sys     0m1.448s
```

Diferencia de tiempo entre ejecución secuencial y concurrente

Ejecución secuencial: 17,456seg

Ejecución concurrente (1 red): 1,816seg

Como se puede ver en los resultados de tiempo obtenidos con las distintas ejecuciones, la versión concurrente supone una gran mejora sobre la versión secuencial.

Mejora = 17,456s/1,816s = 9,612 veces incremento de rendimiento

Diferencia de tiempo entre las ejecuciones concurrentes variando n_reducers

Ejecuciones secuencial y concurrente