

學號:F74074122

姓名:歐禮寬

系級:資訊 111

開發環境:

OS: Windows Subsystem for Linux Ubuntu 18.04.1

CPU: Intel® Core™ i5-8250U CPU @ 1.60GHz

Memory: 8GB

Programming Language: C++ gcc version 7.4.0

程式執行時間:

排序 10GB 約 950~1050 秒吧

程式開發與使用說明:

將一個大檔案(input.txt)切成一個個小檔案(小檔案的排序使用 C++內建的 sort)，最後在用 k-way merge(k 為分割出來的檔案數)合併輸出(result.txt)，t 程式結束前會把切割的小檔案刪除。

編譯: g++ oshw1externalsort.cpp

執行: ./a.out

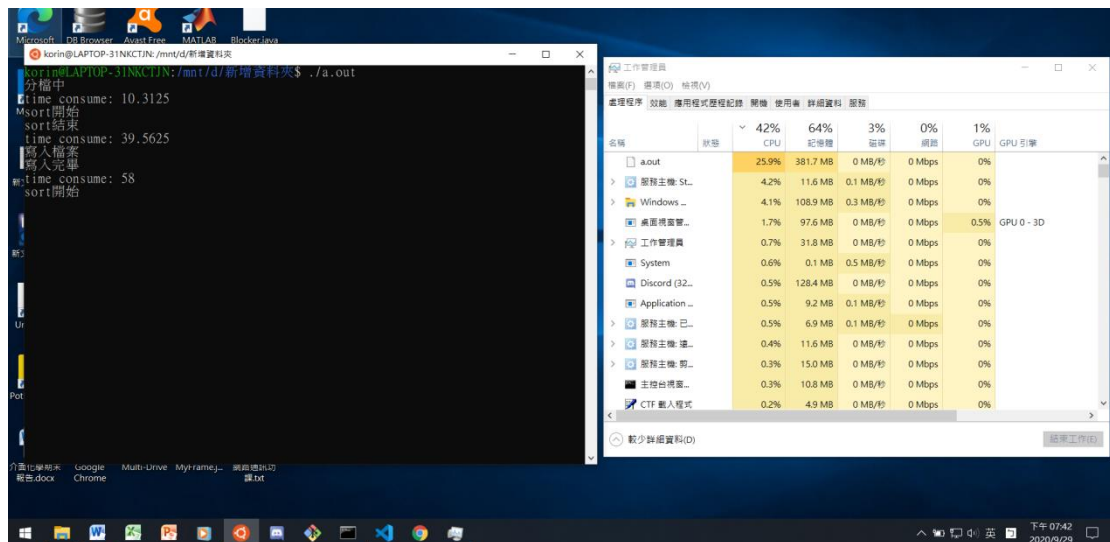
效能分析報告:

優化:

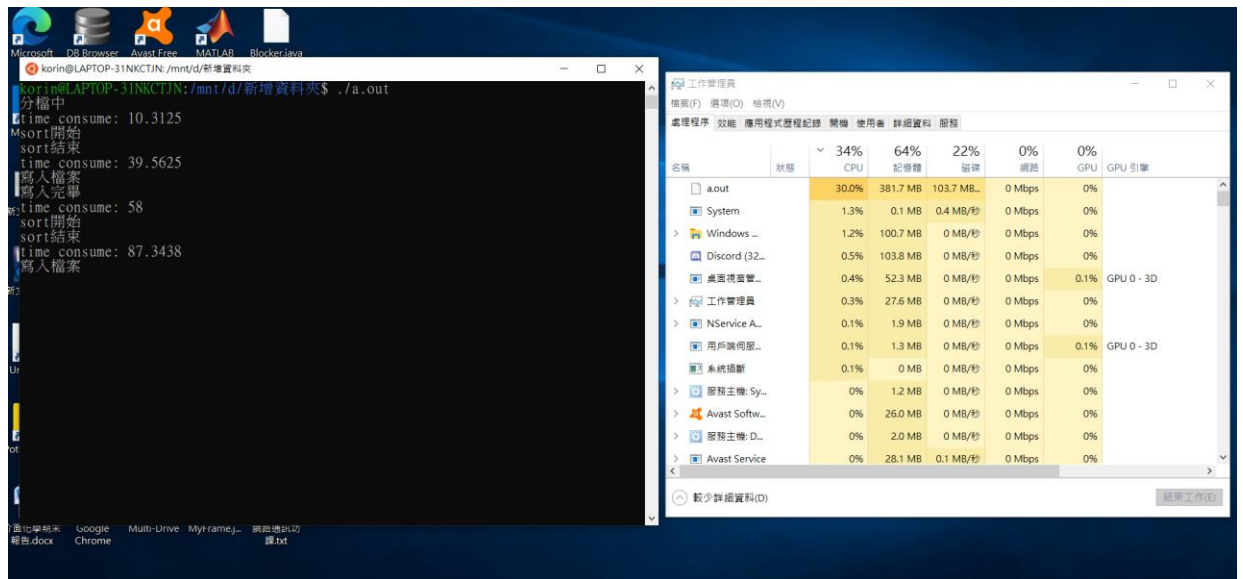
原本我用 C++ 的 ifstream 和 ofstream 來讀寫檔案，結果 I/O 的速度非常緩慢，大概 0.1MB/秒，後來改用 fprintf 及 fscanf 來讀寫檔案，I/O 的數值就增高到大概 100MB/秒。

(1).只開一支程式

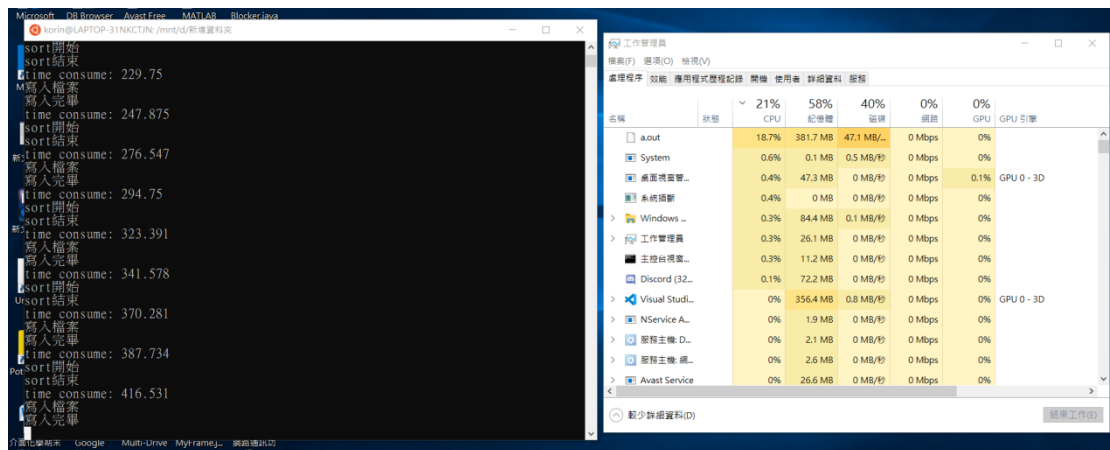
Sort 開始時，用到 C++ 內建 sort 函式排序很大筆的資料，CPU 使用率很高



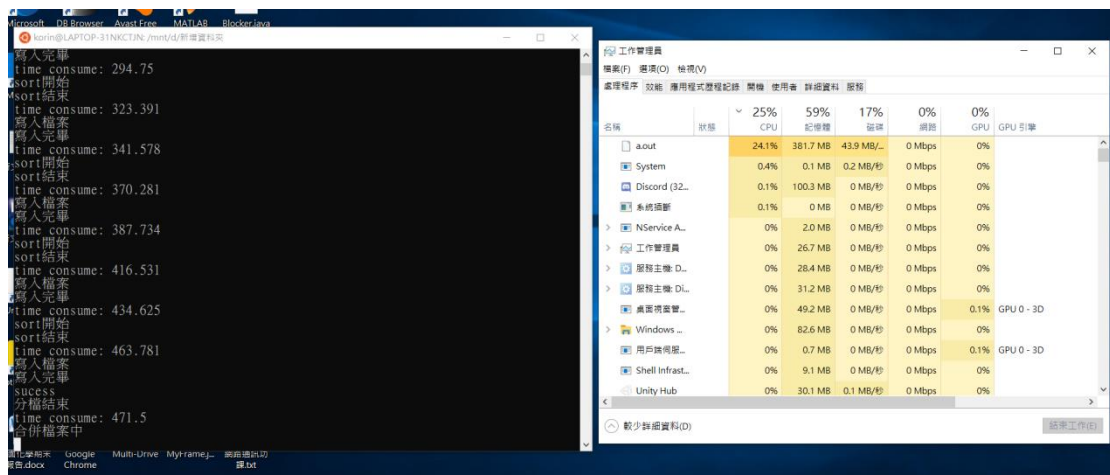
寫入檔案時，進行檔案輸出，這時 CPU 及磁碟使用率很大



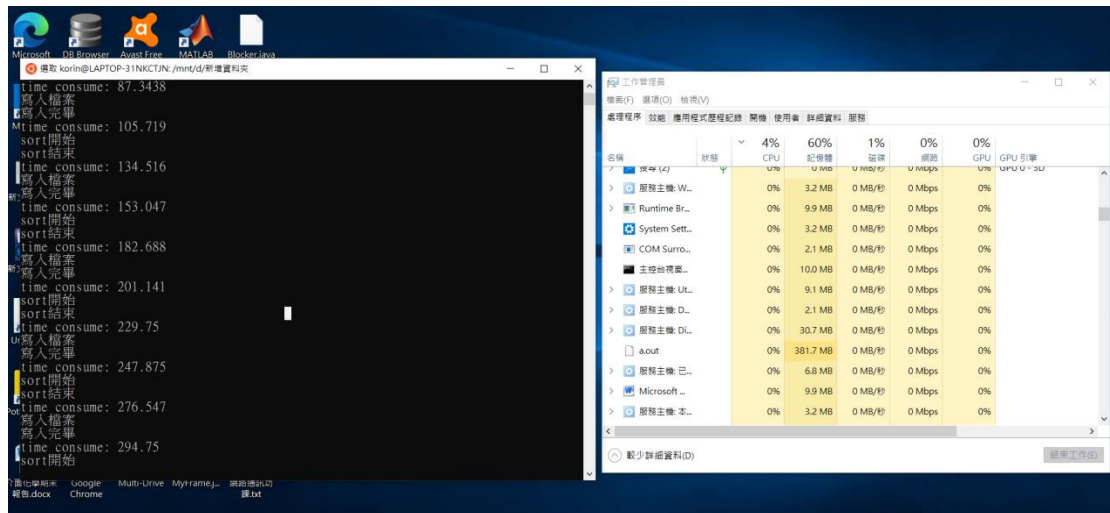
檔案寫入完畢，緊接著又要繼續讀檔，磁碟使用率很高



合併檔案時，用到 sort 及檔案輸入和輸出，CPU 使用率及磁碟使用率都蠻高的

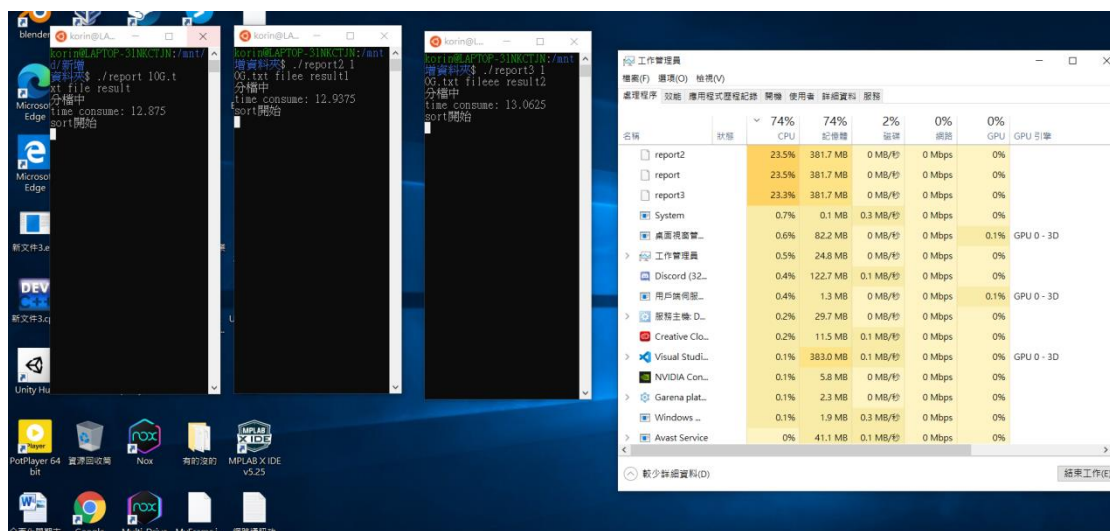


有時 CPU 及磁碟使用率會降到 0，可能是作業系統突然在做某些事情的緣故吧，通常過一陣子就恢復正常了！

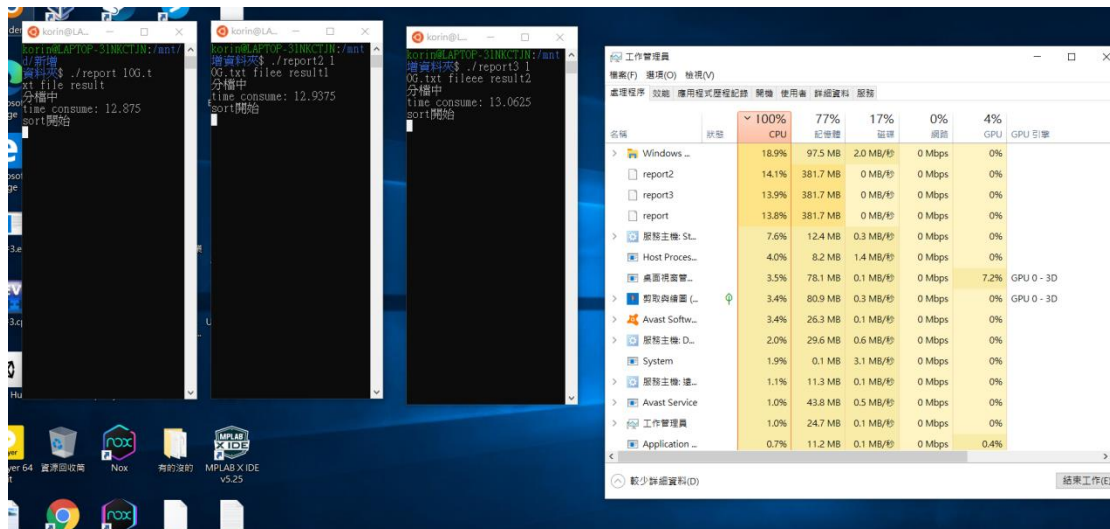


(2.)同時開多支程式

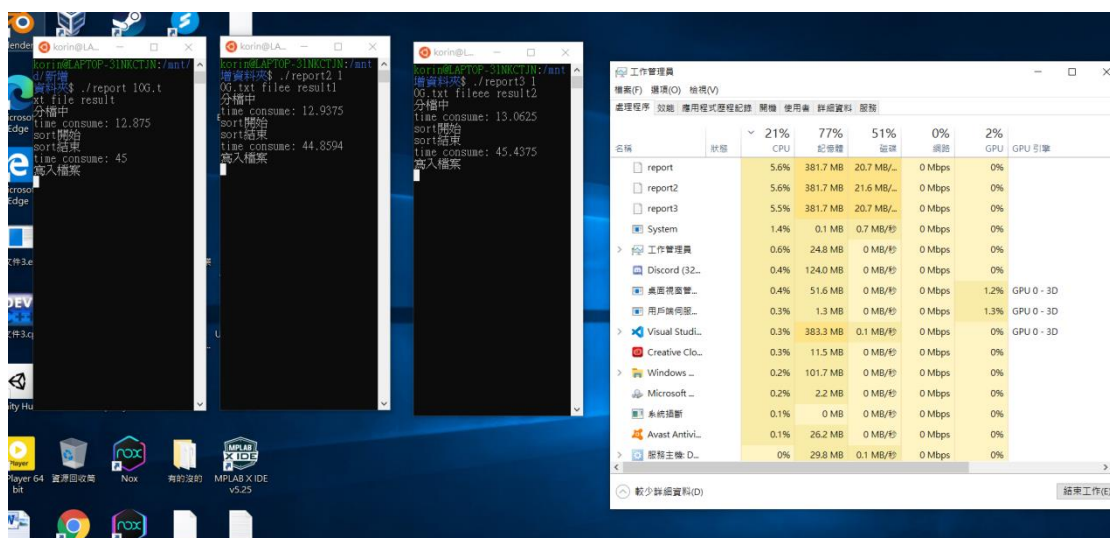
Sort 開始時，CPU 使用率一樣很高，跟只開一支程式差不多



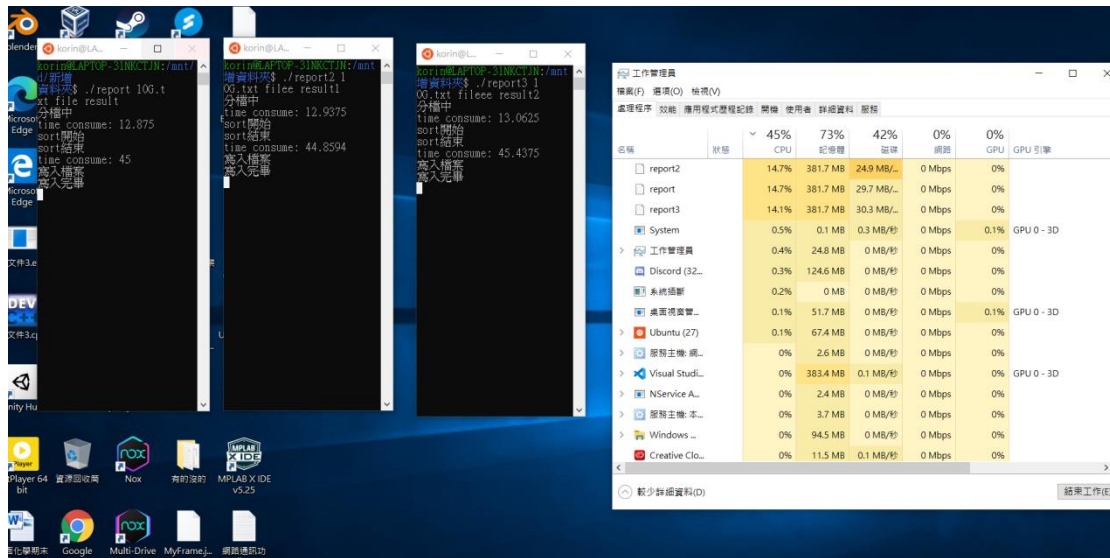
但開了其他東西，CPU 負荷不了便會降低 CPU 使用率



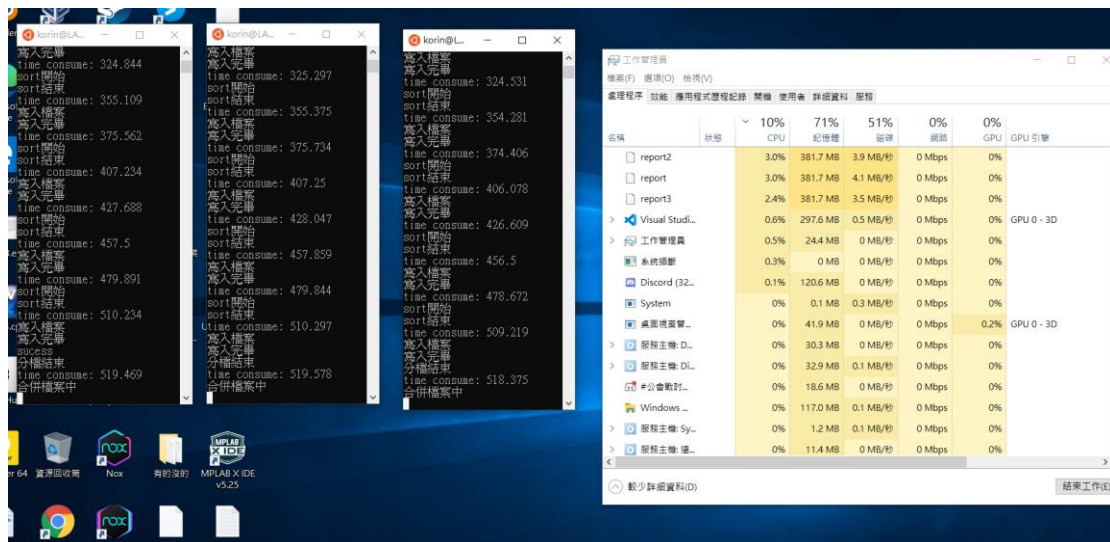
寫入檔案時，比起執行單支程式，CPU 使用較低，且磁碟使用率無法像單支程式那麼高



寫入完畢後，跟只執行單支程式相比，磁碟使用率也無法向單支程式一樣高



合併檔案時，磁碟使用率也比執行單支程式要低許多



總結:

與只執行一支程式相比，同時執行三支程式磁碟使用率(I/O)會變慢非常多，而 CPU 使用率反而不會有顯著的變少，所以原本一支程式 sort 10G 的資料只需九百多秒，要是同時跑 3 支程式，每支程式 sort 完 10G 的資料卻要遠大於九百多秒。而程式所使用的記憶體

大小好像不太會變動，程式執行期間，記憶體大部分時間都是固定的值，不像玩遊戲時，程式使用的記憶體會忽高忽低。