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

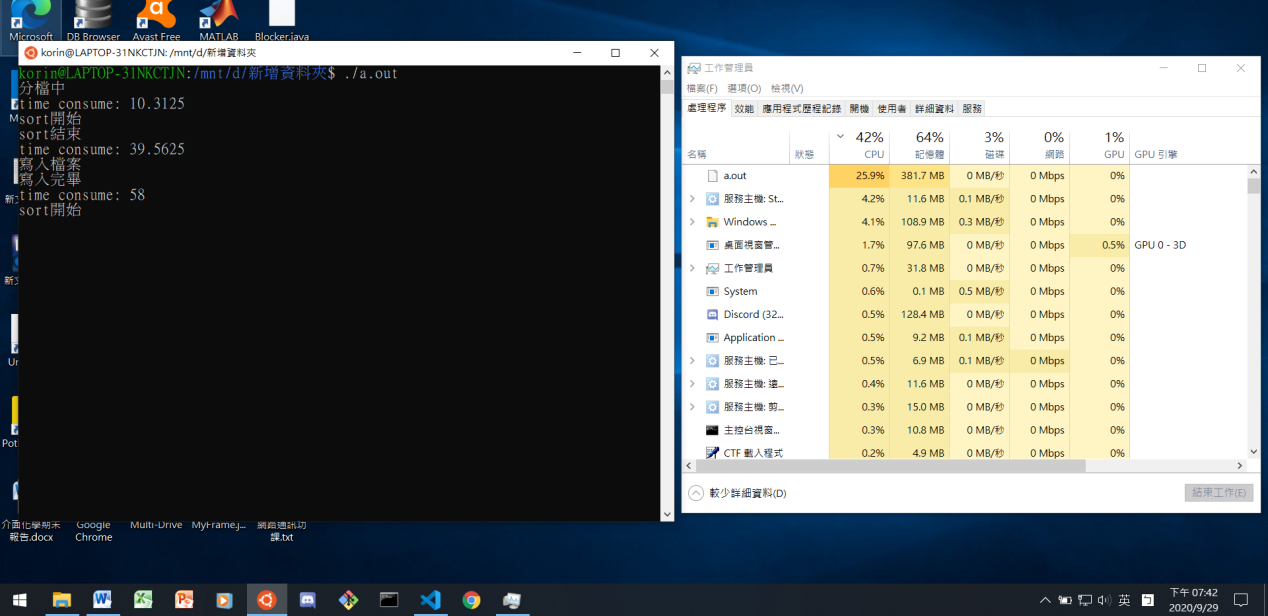
效能分析報告:

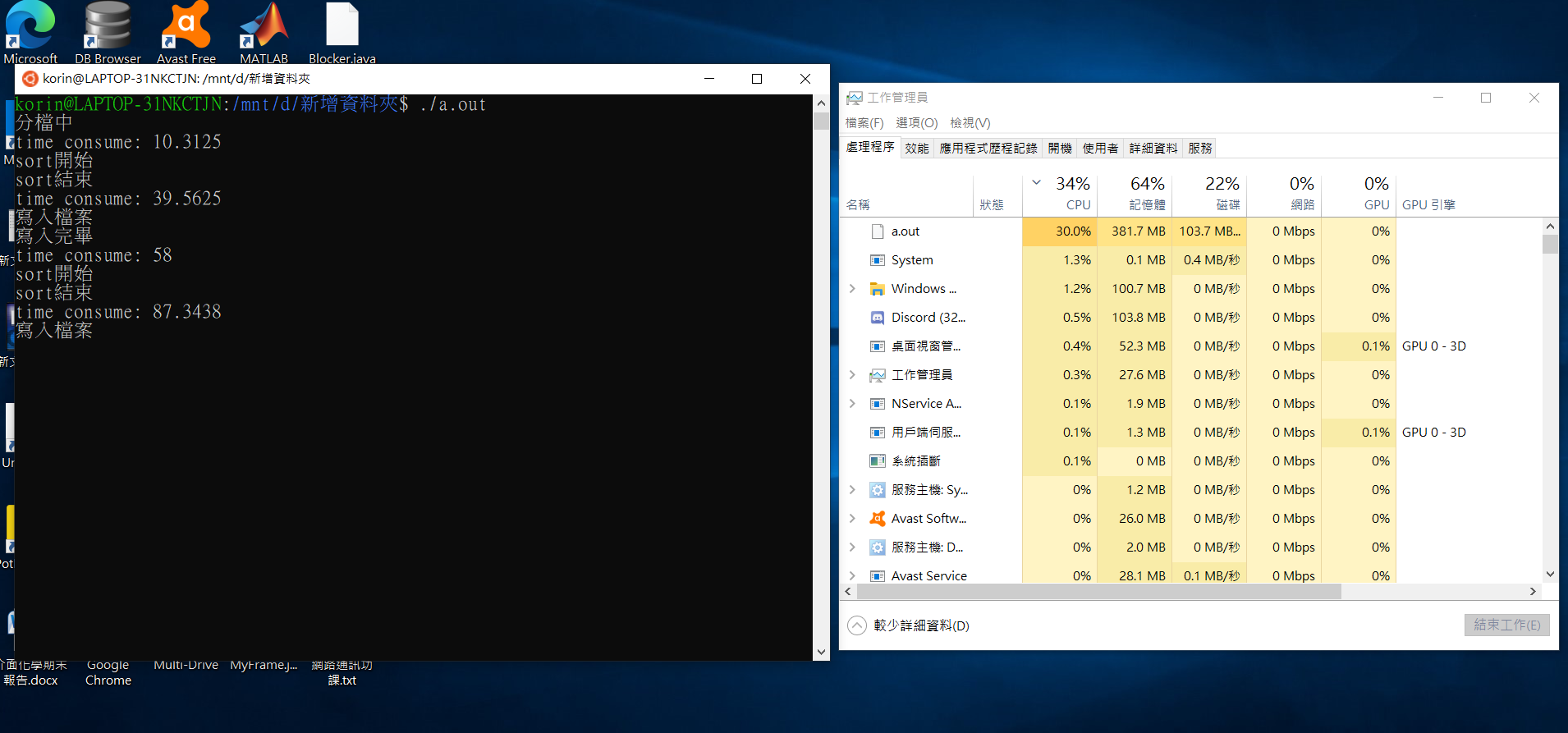
優化:

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的數度就增高到大概100MB/秒。

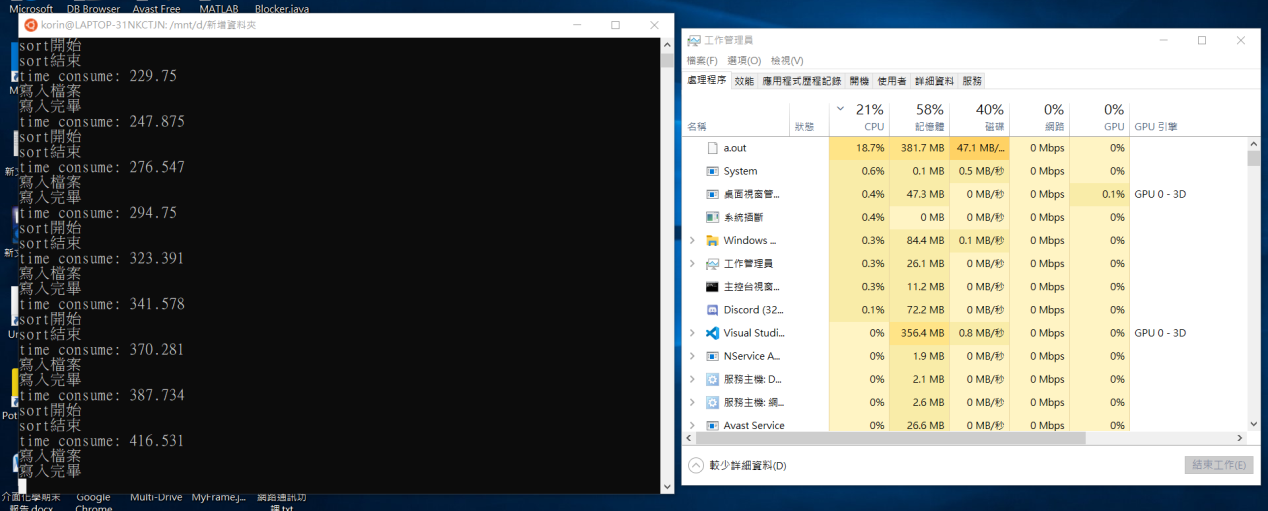
(1).只開一支程式

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

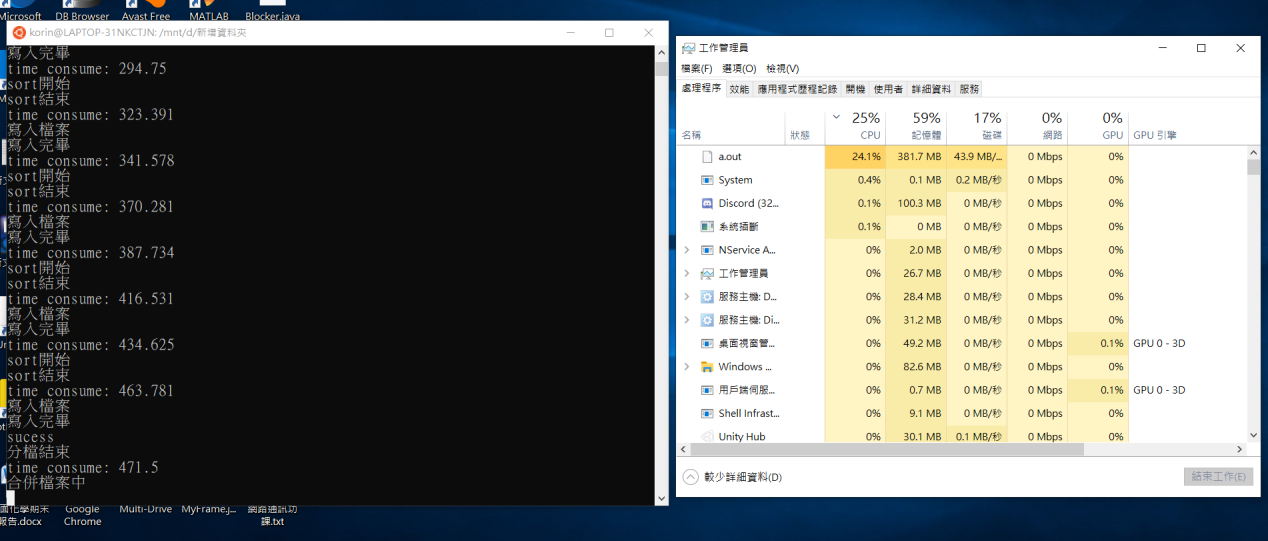


寫入檔案時，進行檔案輸出，這時CPU及磁碟使用率很大 

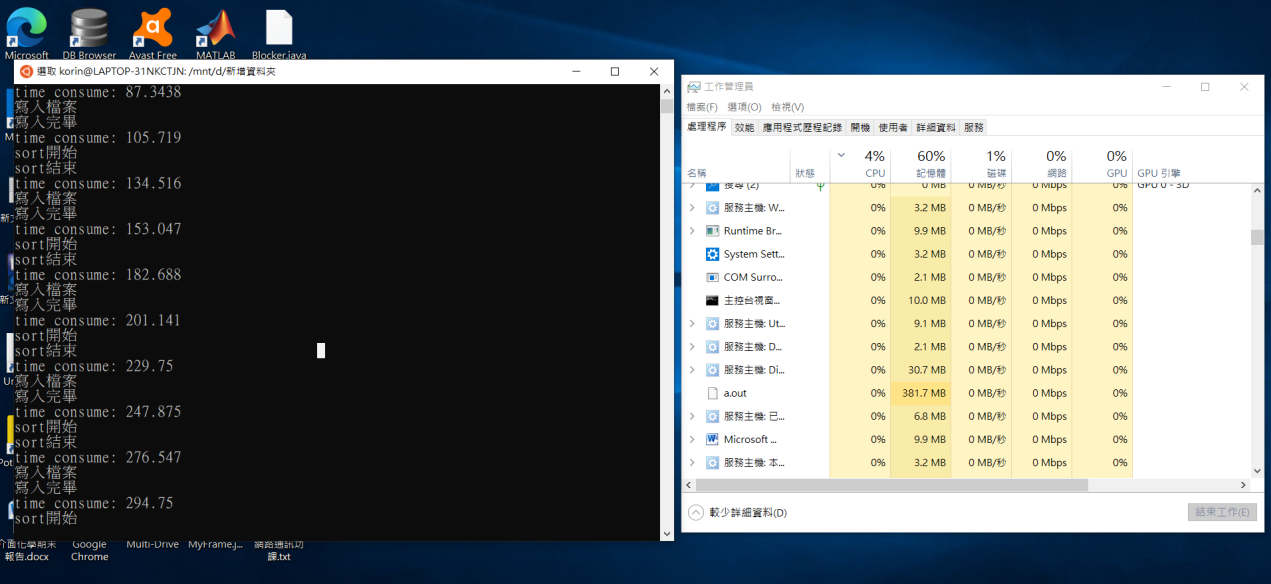
檔案寫入完畢，緊接著又要繼續讀檔，磁碟使用率很高



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

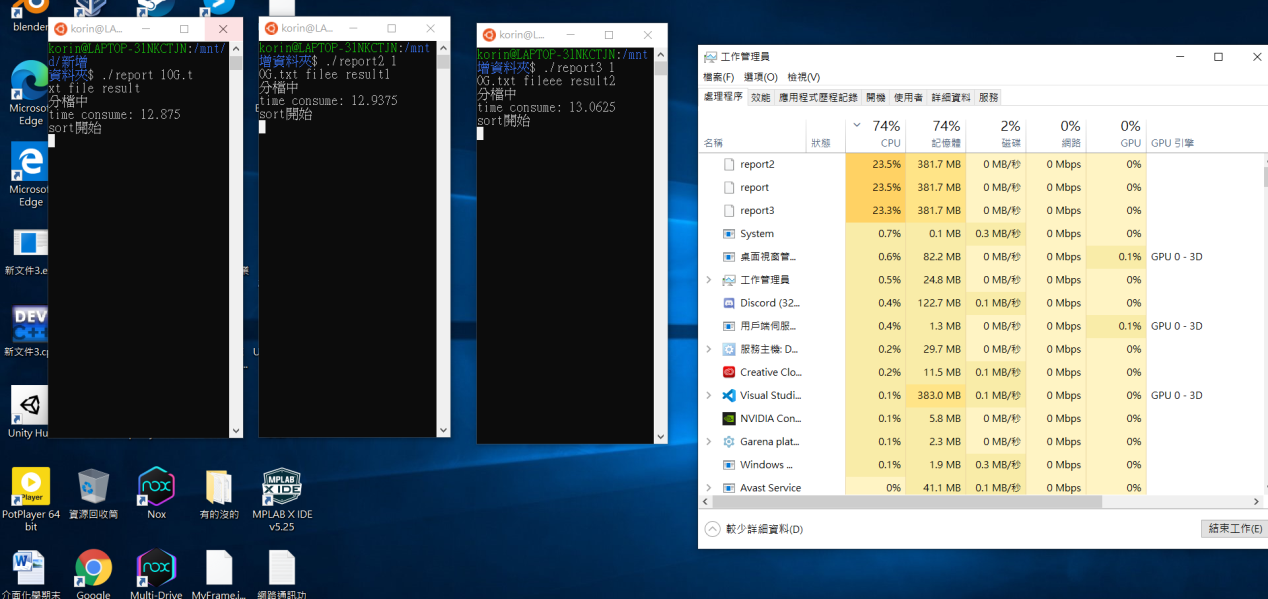


有時CPU及磁碟使用率會降到0，可能是作業系統突然在做某些事情的緣故吧，通常過一陣子就恢復正常了!

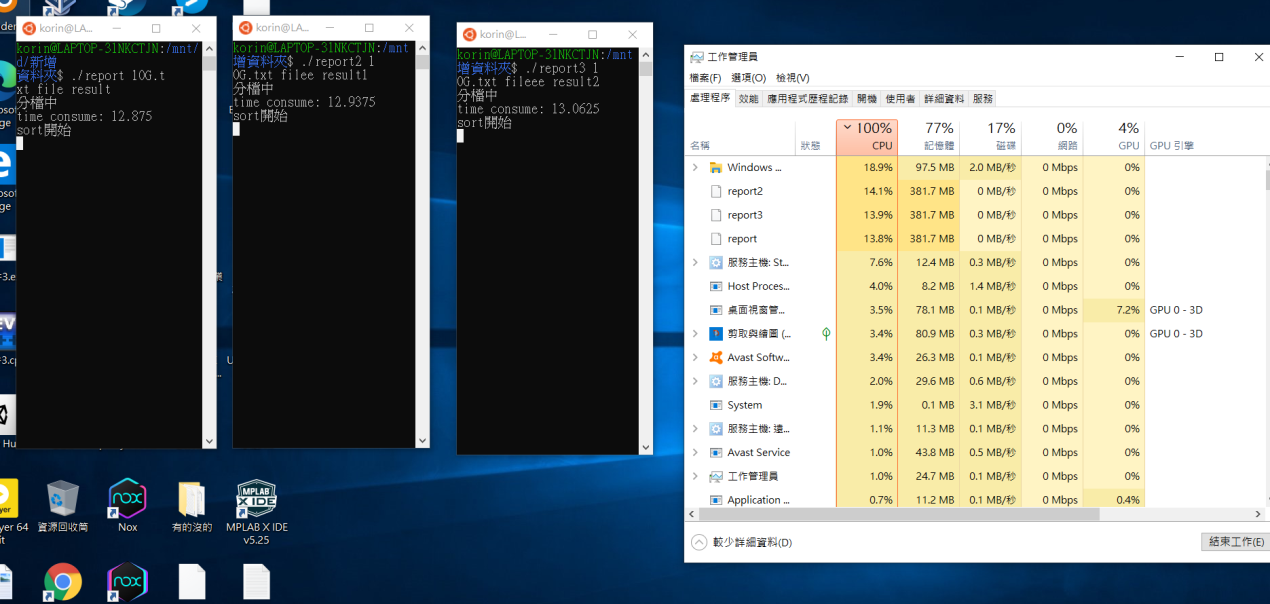


(2.)同時開多支程式

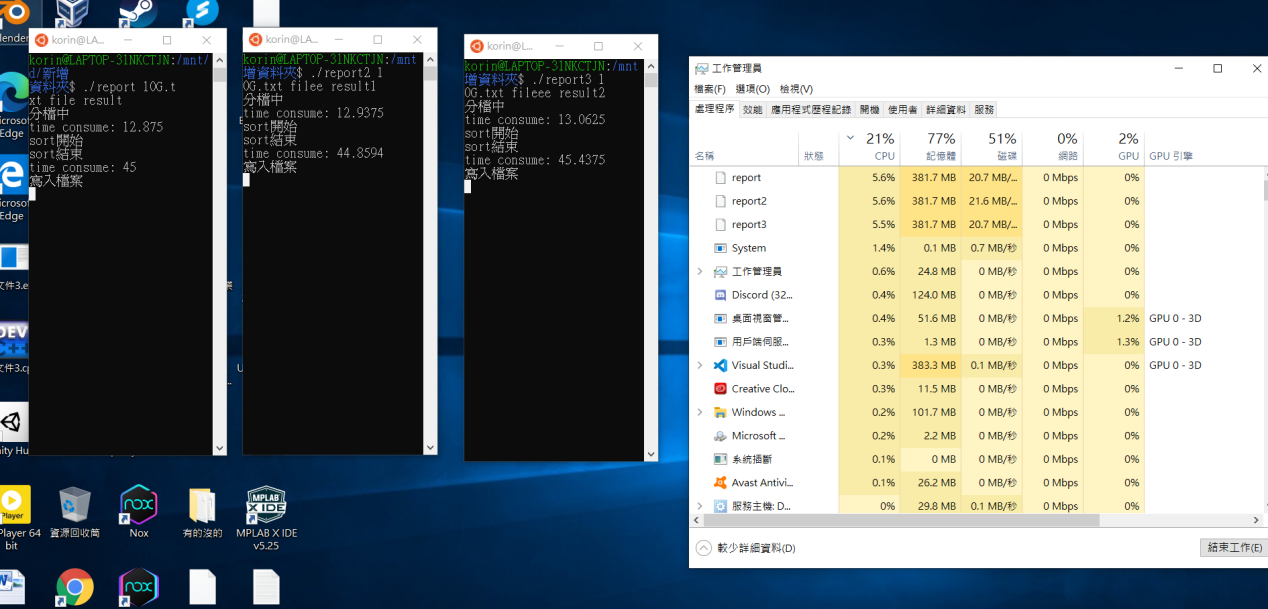
Sort開始時，CPU使用率一樣很高，跟只開一支程式差不多



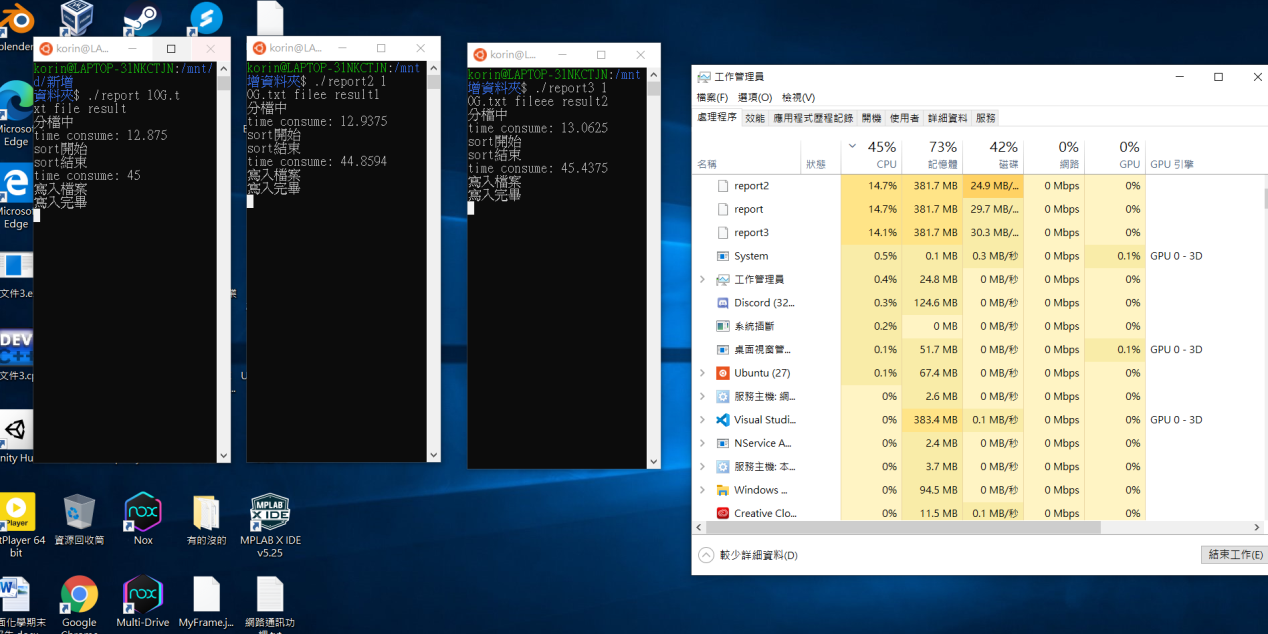
但開了其他東西，CPU負荷不了便會降低CPU使用率



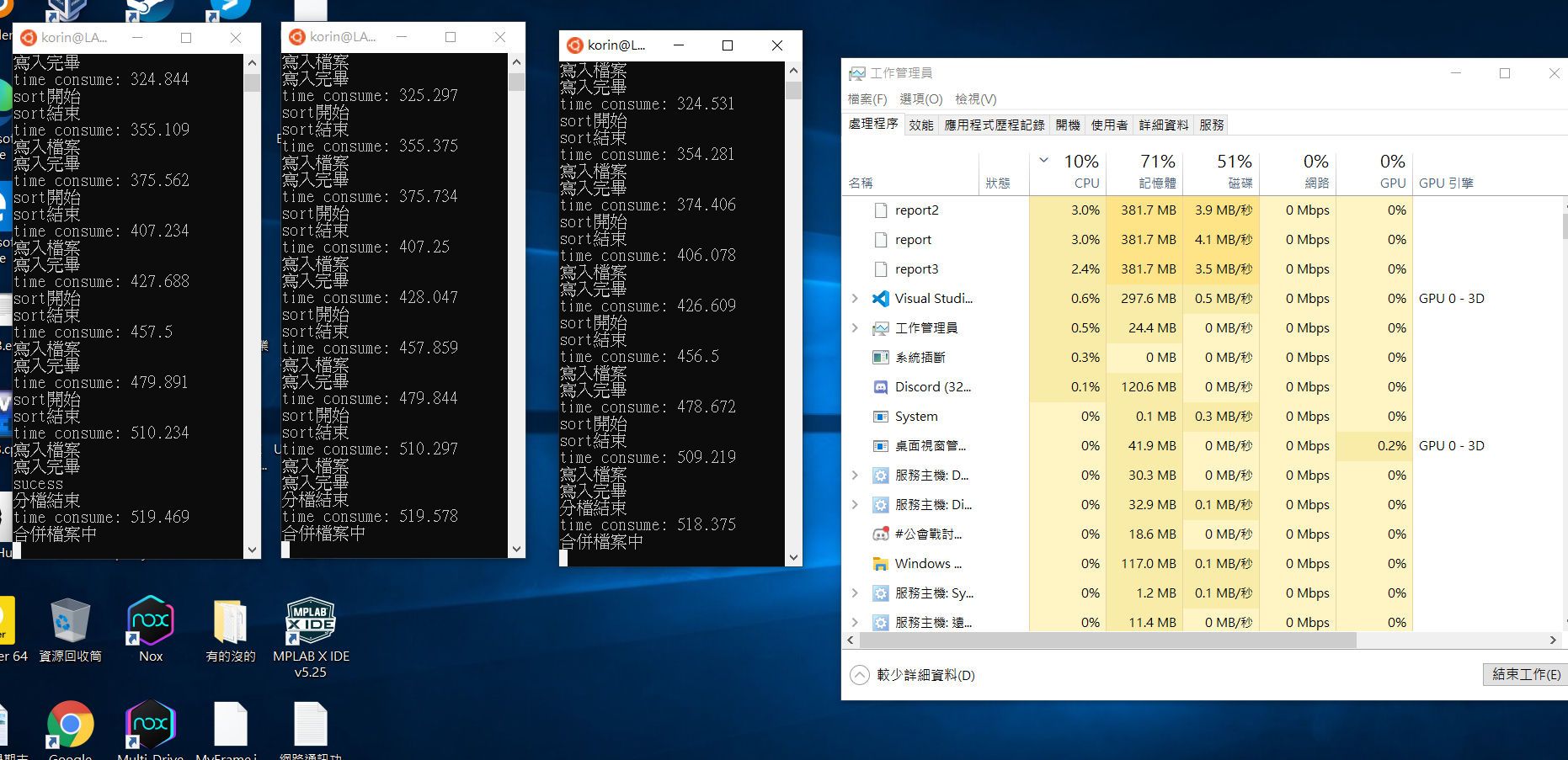
寫入檔案時，比起執行單支程式，CPU使用較低，且磁碟使用率無法像單支程式那麼高



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。