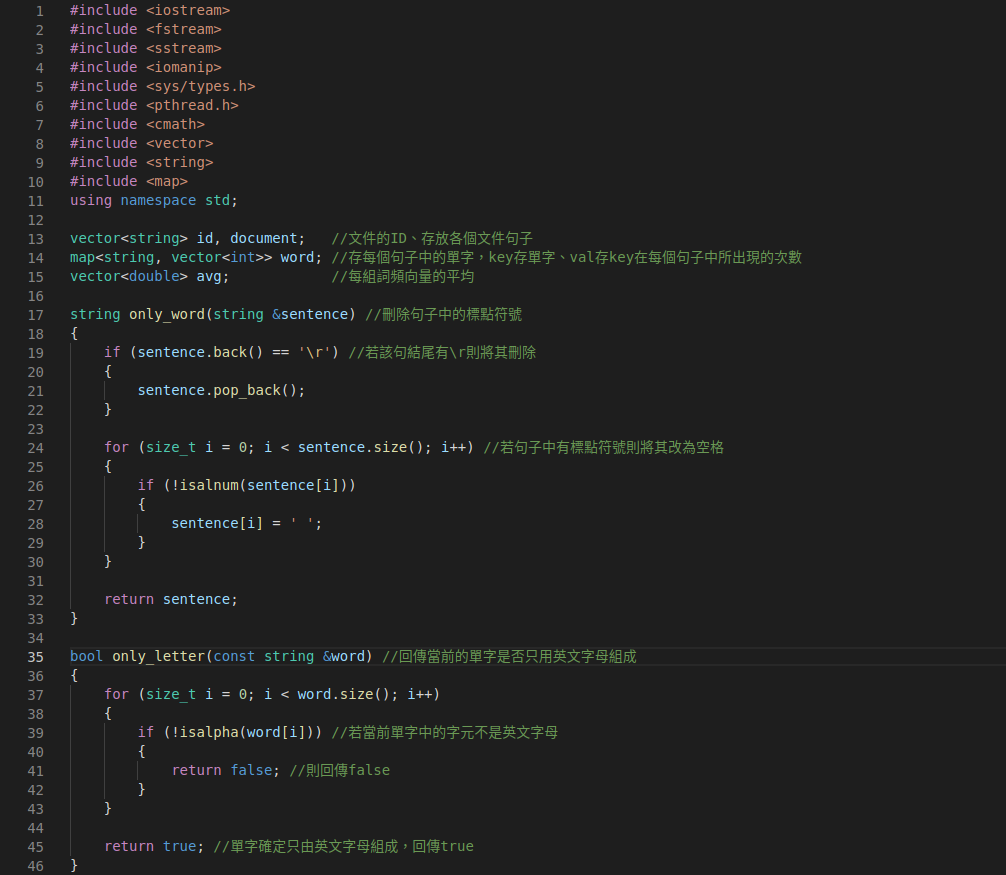
## 1093304作業系統\_hw2 說明報告

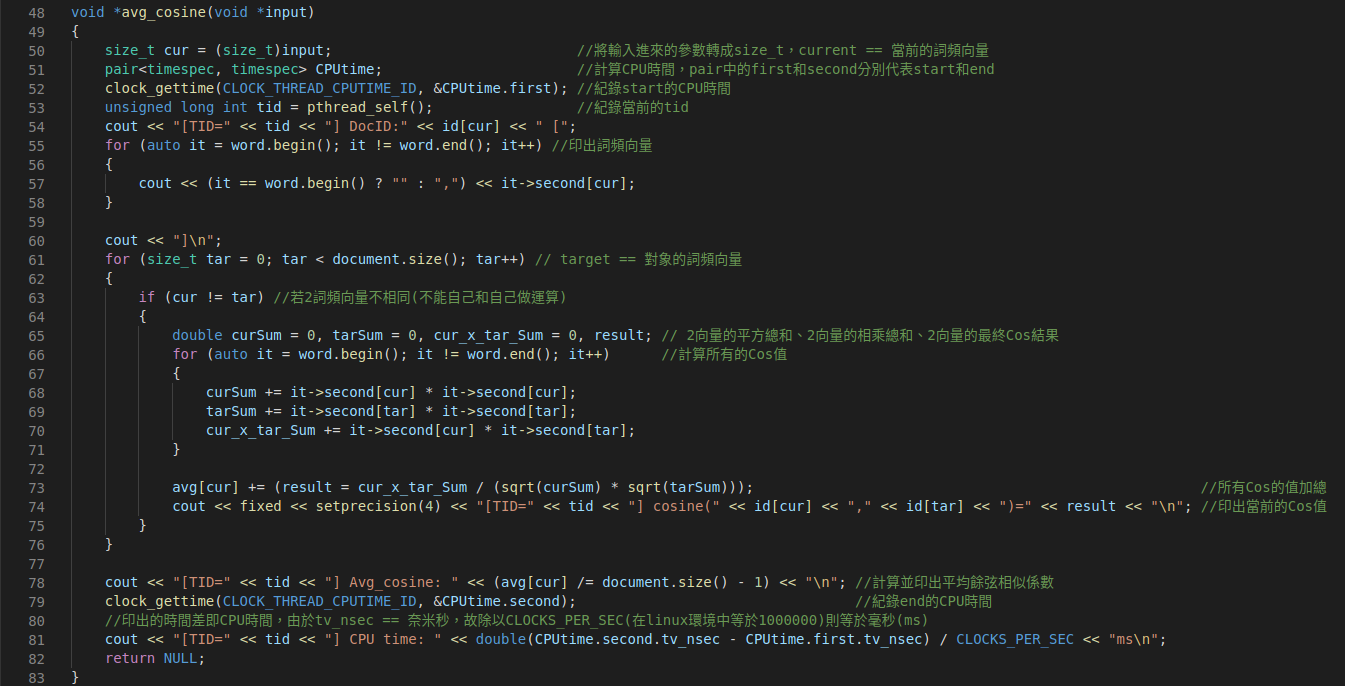
### 設計理念：



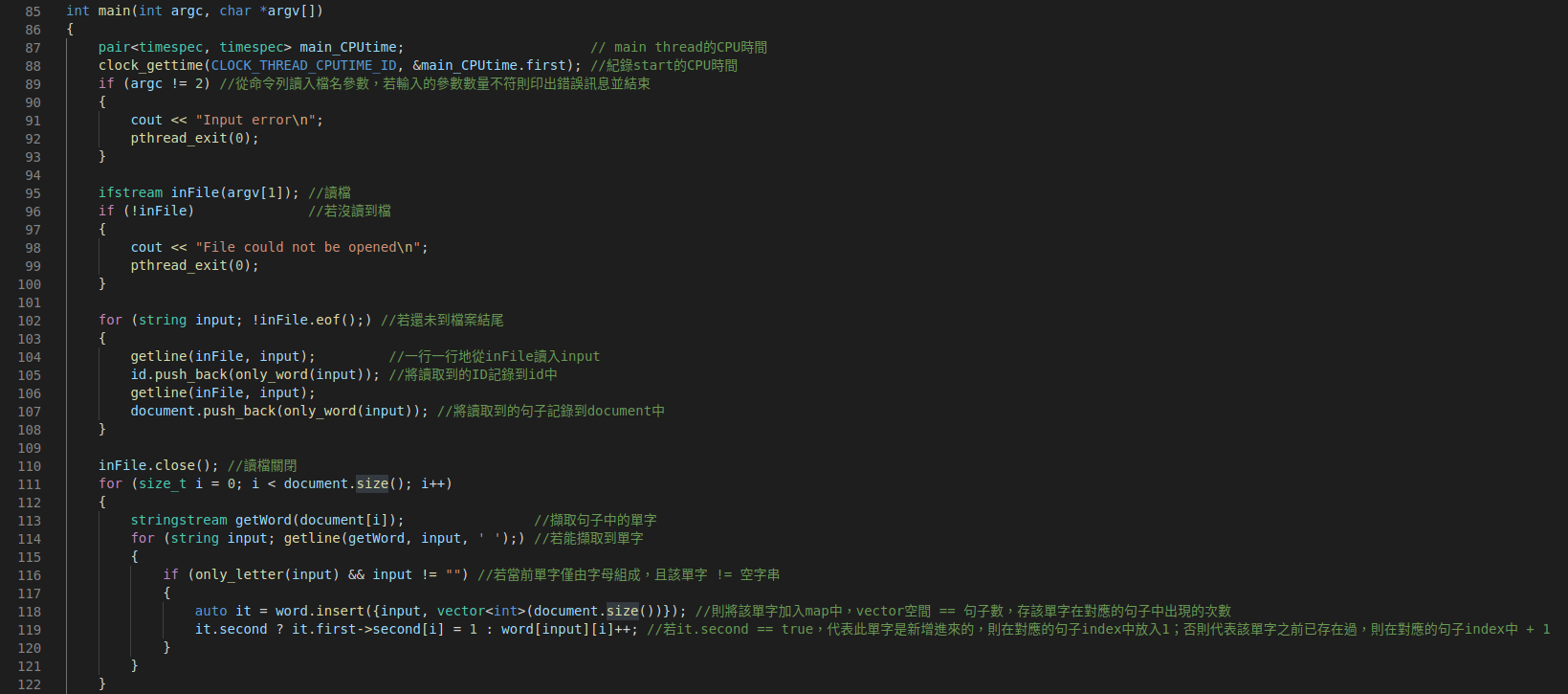
第13～15行：建立存文件id和各文件句子的vector，map存每個單字在每個文件中出現的次數、avg存每個文件和其他所有文件所算出的Cos平均。

第17～33行：對文件做初步處理，主要是將標點符號和\r去掉。

第35～46行：檢查當前擷取的單字是否只用英文字母組成。



第48～83行：計算平均cosine的函式，其中第52行和79行是算出當下的CPU時間，最後在第81行印出2者的時間差就是花費的CPU時間。

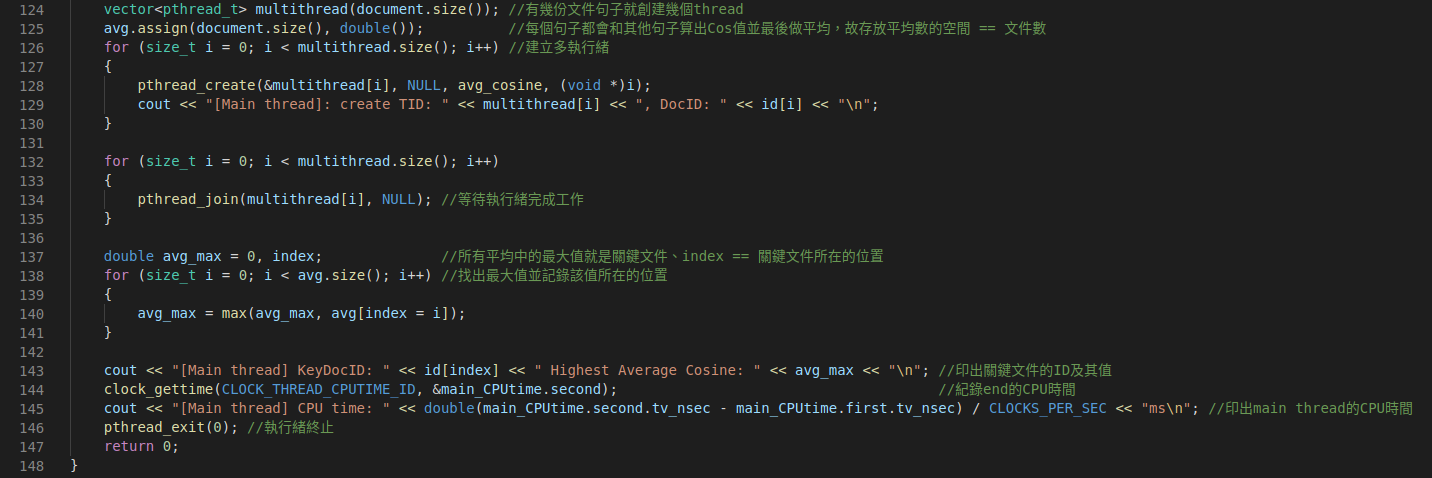


第88行：紀錄main thread當下的CPU時間。

第89～100行：若輸入或讀檔錯誤，則印出錯誤訊息並結束程式。

第102～108行：將讀進來的ID和文件存到對應的vector中。

第111～122行：擷取每份文件中的單字，若單字只由字母組成則放進map中，其中在118行，insert回傳的值是一個pair<iterator, bool>，bool表示當前的insert是否成功。若bool值等於true，則代表原本的map沒有這個單字，所以insert成功，map內的元素會 + 1，並且iterator指向該元素；反之則代表欲insert的key已存在map中，所以insert失敗，其iterator指向指定的key的元素。



第124行：根據文件數創建thread。

第126～130行：建立多執行緒。

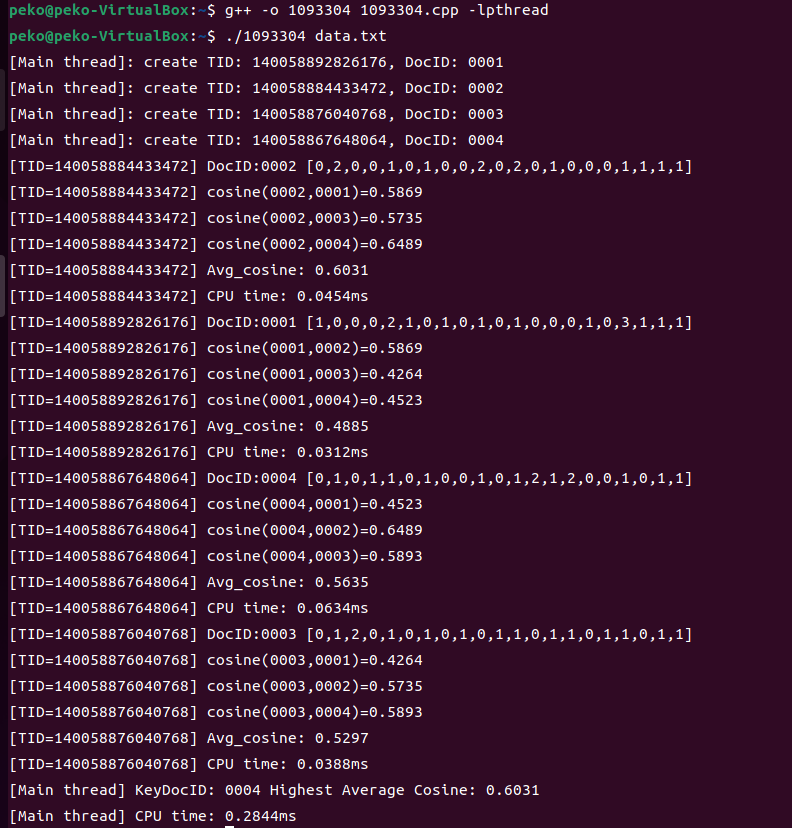
第132～135行：等待執行緒完成工作。

第137～141行：找出所有平均中的最大值。

第144行：紀錄main thread當下的CPU時間。

第145行：2時間相減即為main thread的CPU時間。

### 執行畫面：



開啟終端機，輸入g++ -o 1093304 1093304.cpp -lpthread，再輸入./1093304 +讀檔文件名稱(如圖中的data.txt)後便可執行程式。