CS 307 HW4

Ayşegül Rana Erdemli 26744

The file is read directly from the disk with the C++ method fstream (I used ifstream in this). It takes a long time since accessing to the disk is slow. For each operation we need to access to the disk since we do not have the information in the main memory. Since our file was large, both fstream and fopen was slow at reading. But C function fopen was faster than fstream so that reading a file with fopen is less time consuming.

```
[aysegulrana@flow ~]$ g++ -o cs307hw4.out cs307hw4.cpp -std=c++11
[aysegulrana@flow ~]$ ./cs307hw4.out
the number of occurences of 'a' in the file is 19082160
cpp fstream time: 5791409785 nanoseconds
```

The function fpoen in C programming does a similar thing with fstream in C++.

```
[aysegulrana@flow ~]$ ./cs307hw4.out
the number of occurences of 'a' in the file is 19082160
c fopen time: 1609784696 nanoseconds[aysegulrana@flow ~]$
```

Using memory map, we prevent long read and write operations from the disk and use virtual memory, our program will act like our memory mapped file is in the main memory and access its information from there. This way is the most efficient one since we can access the data like an array since we mapped the file to a process address space. The file can be accessed really fast and that was the case in my program too. It took less than 1 nanoseconds. The disadvantage of memory mapping is consuming the RAM but they are very fficient since we have a large file to read. A disadvantage of it is more likely to cause a bus error than reading file from the disk.

```
[aysegulrana@flow ~]$ ./cs307hw4mmap.out
the number of occurences of 'a' in the file is 19082160
c mmap time: 0 nanoseconds[aysegulrana@flow ~]$
```