Doğan Can Hasanoğlu
26809

CS307 HW4 REPORT

METHODS

- **fstream (C++)**
- **fopen (C)**
- **memory mapping - mmap (C)**

Total occurrences of character ‘a’ : 19082160

Average time taken by each method:
fstream = ~ 6,5 seconds (did not use getline())
fopen = ~ 3,5 seconds
memory mapping = ~ 1,5 seconds

First of all as you can see memory mapping is the fastest one by far. I will discuss it later on but now I’d like to discuss **fstream** vs. **fopen**.

Fstream method in C++ is slower than the fopen method in C. Main reason for that is that fstream makes more function calls for reading while fopen returns the bytes read. Besides fstream is more modular and open for manipulation whereas fopen is simple, easier and replaceable. However there is more chance of encountering errors when something is done wrong while using fopen. Advantage of fstream here is fstream is more secure and safe compared to fopen. Last but not least fstream allows you to basically “stream” anything.
In conclusion fstream is slower but more modular, secure and full-scalable. However if you only want to read from a file in a fast manner the way should be fopen in terms of speed and simplicity.

Now more to discuss about memory mapping since it is the significant one here.
Memory mapping utilizes virtual memory. With using memory mapping, mmap() function maps the file to a process address space. With that way, it is known that if we want to make read or write operations on that mapped file, we can reach that file in a more efficient and faster way. Memory mapping is based on mapping a large part of the file from the secondary storage (HDD or SSD) to the memory, and by doing this the physically mapped address becomes directly accessible by the program. This technique brings out higher speed in terms of reading and writing operations on the file. However there are a few disadvantages which are that error handling is difficult, complicated in the system part of the computer (OS) and memory mapping must be dealt within a specific approach -not very user friendly.

In conclusion I’d prefer memory mapping, since it is the fastest way and directly communicates with the OS and utilizes virtual memory. Also fopen is a simple choice if you want to go with a more simple - user friendly- way. Fstream would not be an optimal option if you are not willing to use specifically C++ language.

```
[hasanogludogan@flow ~]$ ./fstream.out
Number of occurences of character 'a' :
19082160

Time taken: 6.64s
[hasanogludogan@flow ~]$ ./fopen.out
Number of occurences of character 'a' :
19082160

Time taken: 3.58s
[hasanogludogan@flow ~]$ ./mmap.out
Number of occurences of character 'a' :
19082160

Time taken: 1.40s
[hasanogludogan@flow ~]$
```