Exercise for _Database System Concepts for Non-Computer Scientist_ im WiSe 18/19

Alexander van Renen (renen@in.tum.de)

http://db.in.tum.de/teaching/ws1819/DBSandere/?lang=en

Sheet 11

Exercise 1

Consider the _Students_ relation in our university schema:

a) Estimate the size (in byte) for an average row of the _Students_ table and then calculate the size of the entire table, assuming there are 1 million students.

b) Calculate the time it would take read the entire table using sequential and random access using an HDD. The seek time is 4 milli seconds, the time for the rotation is 2 milli seconds, and the transfer time is 0.1 milli seconds for a 4 KB block of data (You can ignore the track-to-track-seek-time). You always have to transfer an entire 4KB page, even if you are only reading a single student tuple (in the random access case).

c) Try out the same calculation for an SSD. Here we do not have any seek or rotation time. In addition, the transfer time is as low as 10 micro seconds for reading a 4 KB page.

[Bonus] This exercise is very optional and only if you want to test out your programming skills. Try to write a small C or C++ program to measure the performance difference between random and sequential read of the storage devices on your machine. For this simply create a big file (around 1GB) by writing a bunch of data to your SSD or HDD (you can use malloc to obtain 1GB of memory and then initialize it to any random number). Next, read this data back in: once sequentially from begin to end and once in a random order. Hint: You can use fopen and fclose to open and close a file. And then the fread and fwrite functions to read and write data. For moving around in a file there is the fseek function.