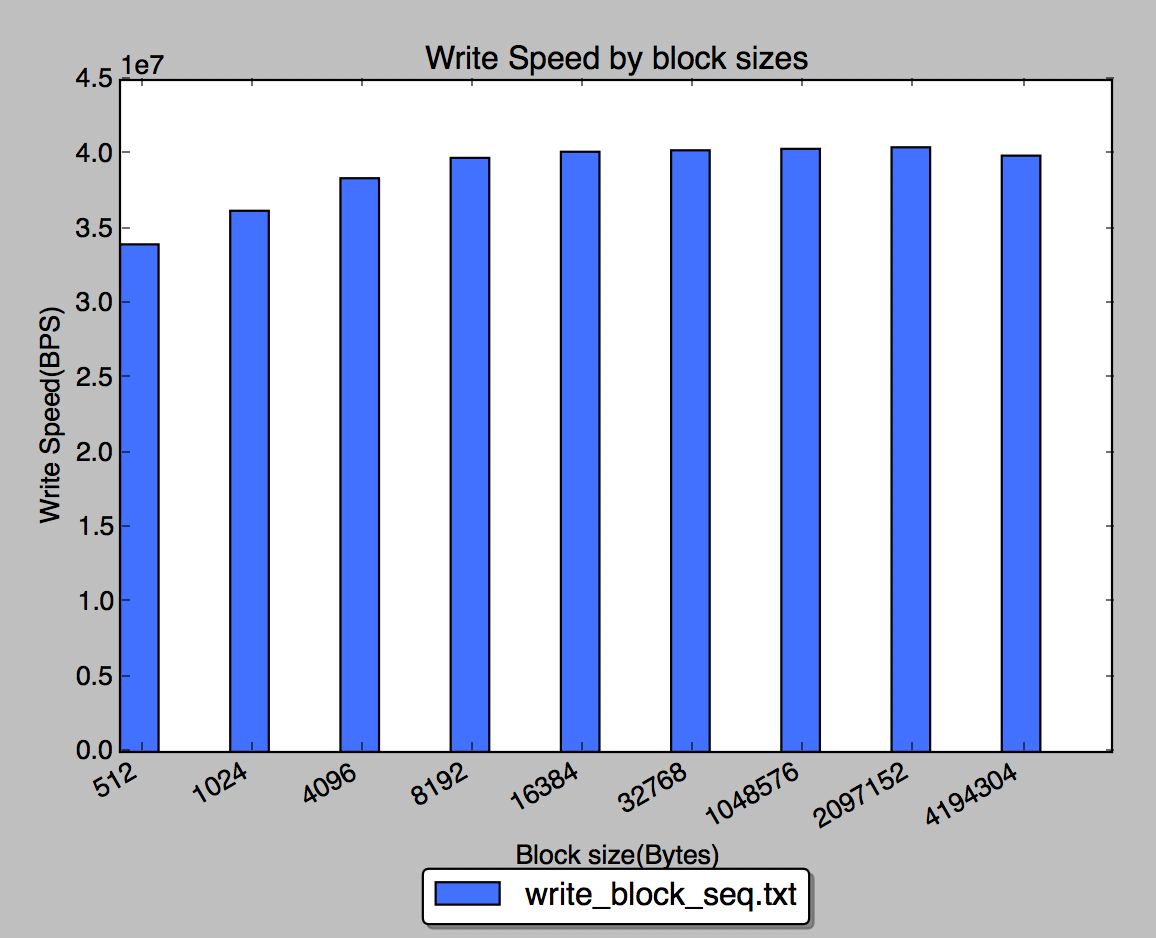


Experiment 1

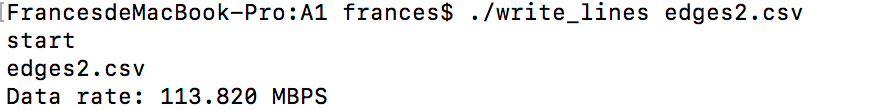
my mac os system block size 4096 bytes



optimal block size to my experiment is 1MB (2MB?

Does it correspond to the system disk block size? No

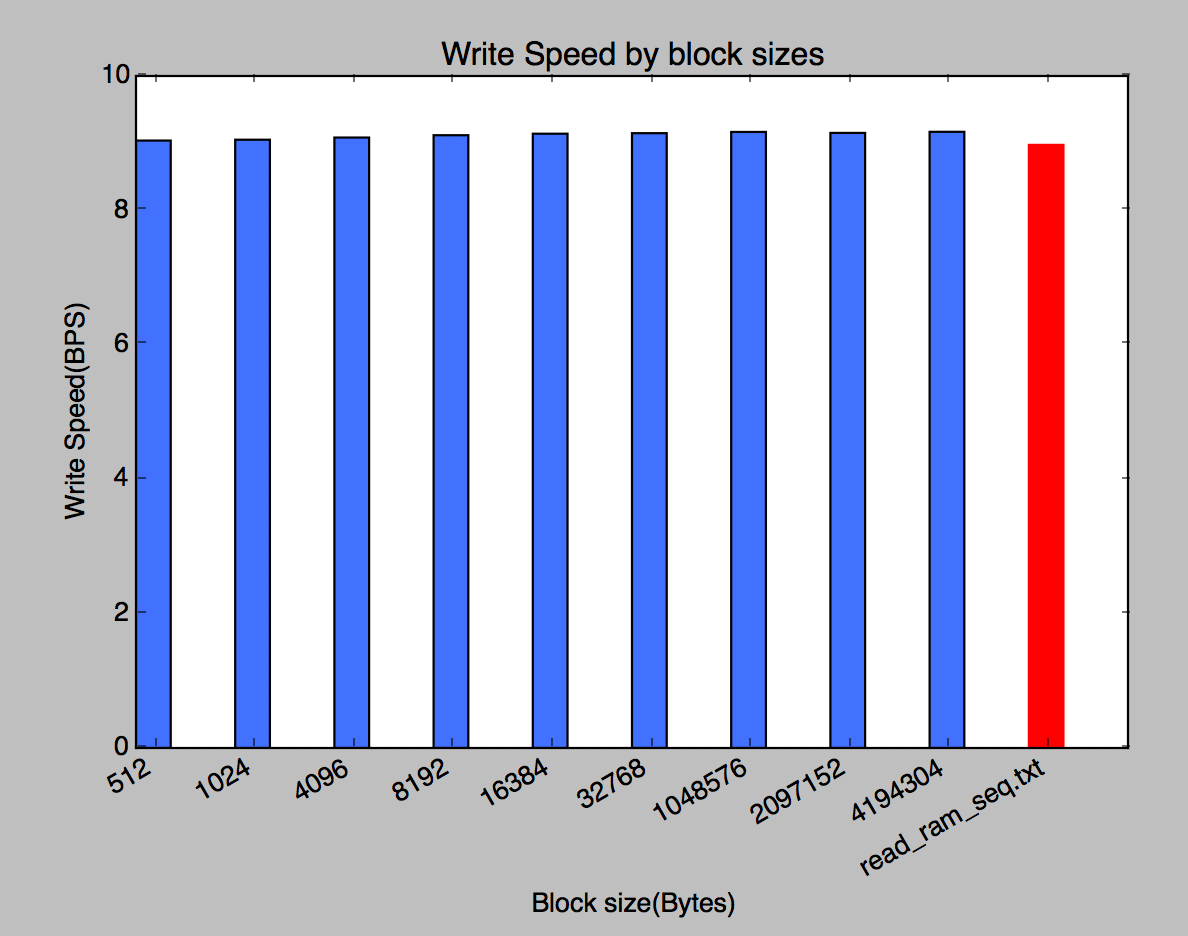
Is there a block size when further increase does not contribute to better performance? 1MB



The speed for writing in lines is 113.820MBPS. It is faster than writing in Blocks. This shows writing in lines is more efficient than writing in blocks.

From the lecture, we learned that using block to write data to disk can reduce disk I/O and it will be more efficient. In our case, the results we got is the opposite way. We thought there might be serval reasons that cause this. First, writing in blocks we need convert each line into record and then push all the records into dish when buffer is full. The converting process may take some time. Second, OS might still help buffering. Writing in lines seems like read from original CSV one line and writes that line to new CSV, but OS can buffer those lines and write to new CSV file at once.

Experiment 2



Q: What is the ratio of sequential read rate for secondary storage and for RAM? Does it correspond to the ratio discussed in class? If not, what do you think is the reason?

A: The ratio discussed in class is around 10^8.8/Sec (sequential RAM) and around 10^7.7/Sec (sequential SSD). But the ratio we got from the data is different, around 10^9/Sec(sequential SSD) and 10^8.7/Sec (sequential RAM). Read speed from RAM is always slower than SSD. The reason we thought is that our laptop RAM is also occupied with other applications when we run the code.

Experiment 3

