Course Project #3: SSD Performance Profiling
Due date: Oct. 25

This project helps you to gain first-hands experience on profiling the performance of modern SSDs (assuming the SSD in your computer is modern enough). The task is simple: Use the Flexible IO tester (FIO), which is available at https://github.com/axboe/fio and may already be included in the OS on your machine, to profile the performance of your SSD. Its man page is https://linux.die.net/man/1/fio. FIO is a storage device testing tool widely used in the industry. Like Project 1, you should design a set of experiments to measure the SSD performance (latency and throughput) under different combinations of the following parameters: (1) data access size (e.g., 4KB/16KB/32KB/128KB), (2) read vs. write intensity ratio (e.g., read-only, write-only, 50%:50% and 70%:30% read vs. write), and (3) I/O queue depth (e.g., 0~1024). Note that throughput is typically represented in terms of IOPS (IO per second) for small access size (e.g., 64KB and below), and represented in terms of MB/s for large access size (e.g., 128KB and above).

WARNING: FIO may overwrite the entire drive partition, so you may want to create an empty partition on your SSD just for FIO testing. Carelessly running FIO on your existing drive partition **will destroy all your data**!

Again, you should observe a clear trade-off between access latency and throughput (as revealed by queueing theory discussed in class): As you increase the storage access queue depth (hence increase data access workload stress), SSD will achieve higher resource utilization and hence higher throughput, but meanwhile the latency of each data access request will be longer.

The specification of Intel Data Center NVMe SSD D7-P5600 (1.6TB) lists a random write-only 4KB IOPS of 130K. Compare your results with this Intel enterprise-grade SSD, and try to explain any unexpected observation (e.g., your client-grade SSD may show higher IOPS than such expensive enterprise-grade SSD, why?). Explain your reasoning in the report on Github.

Post all your results and analysis on Github.