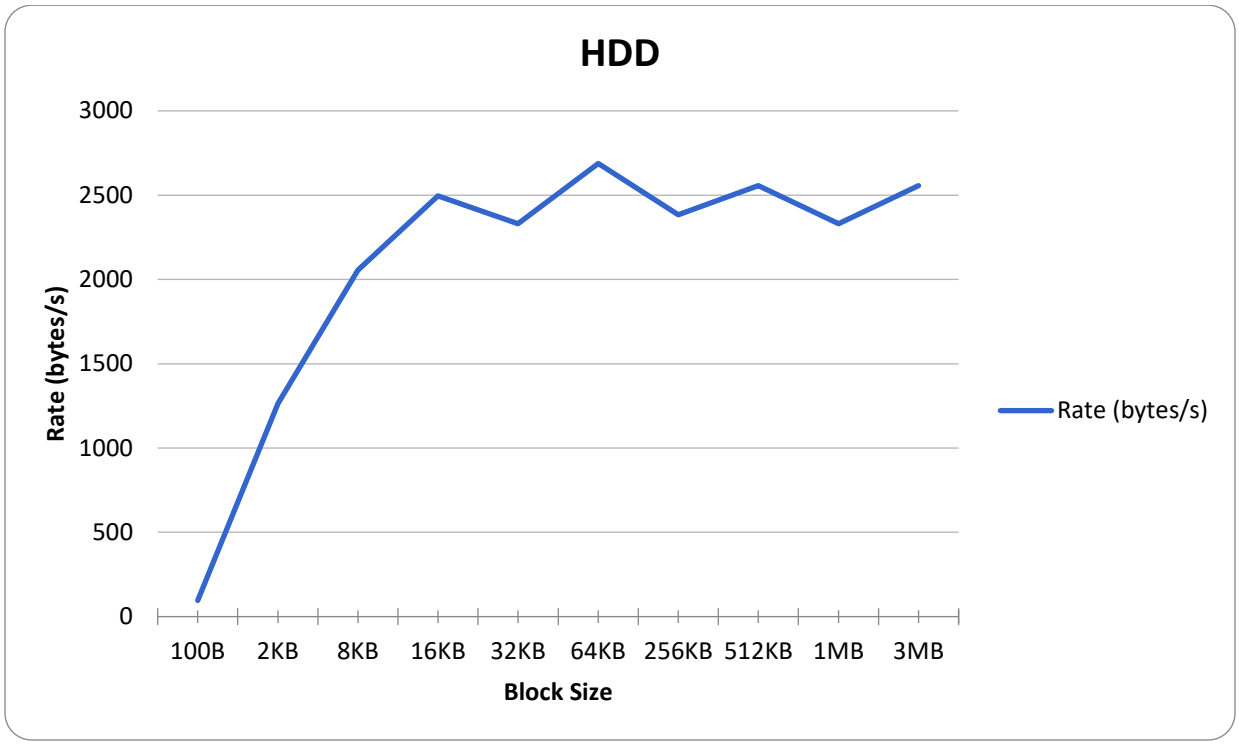
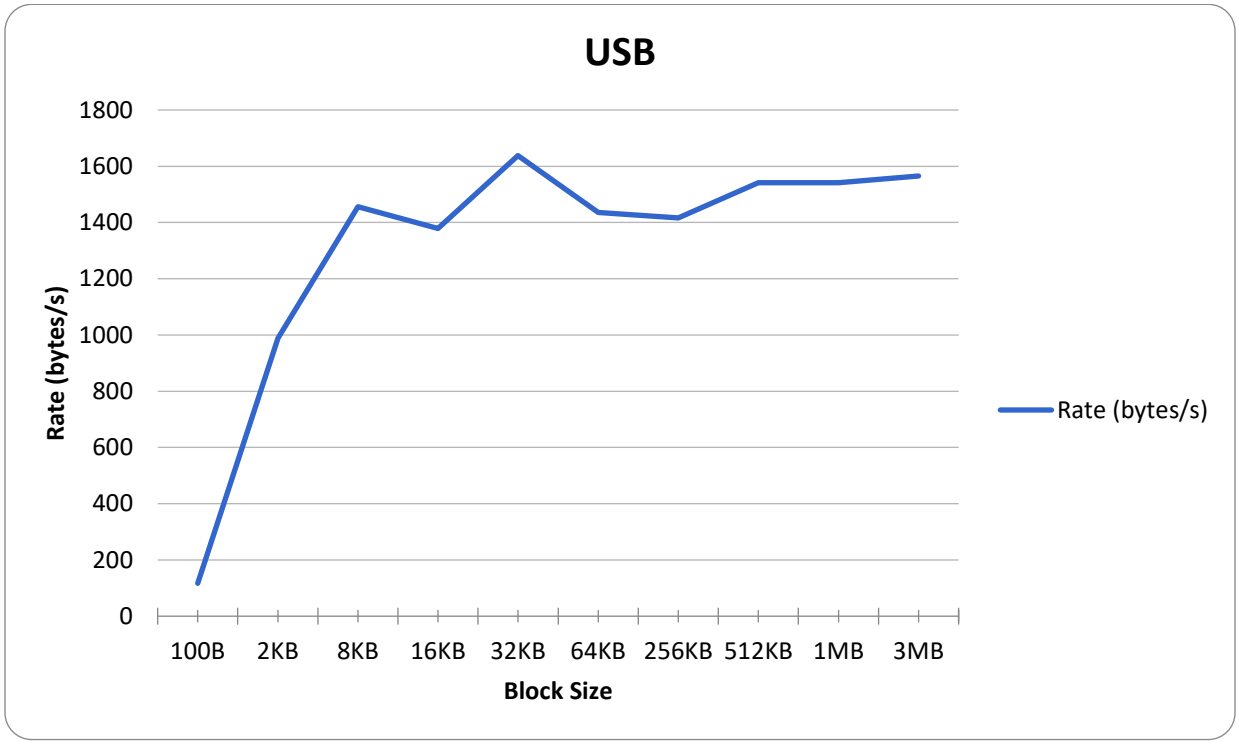
create_random_file





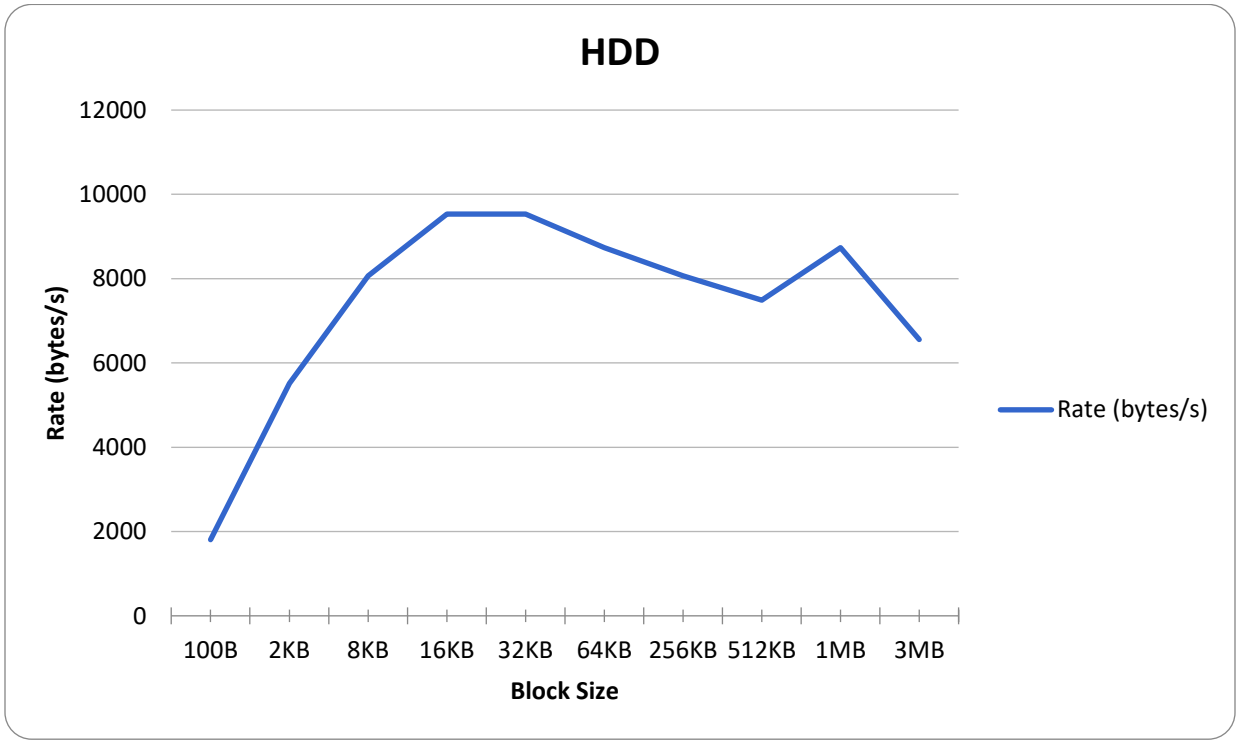
Observations:

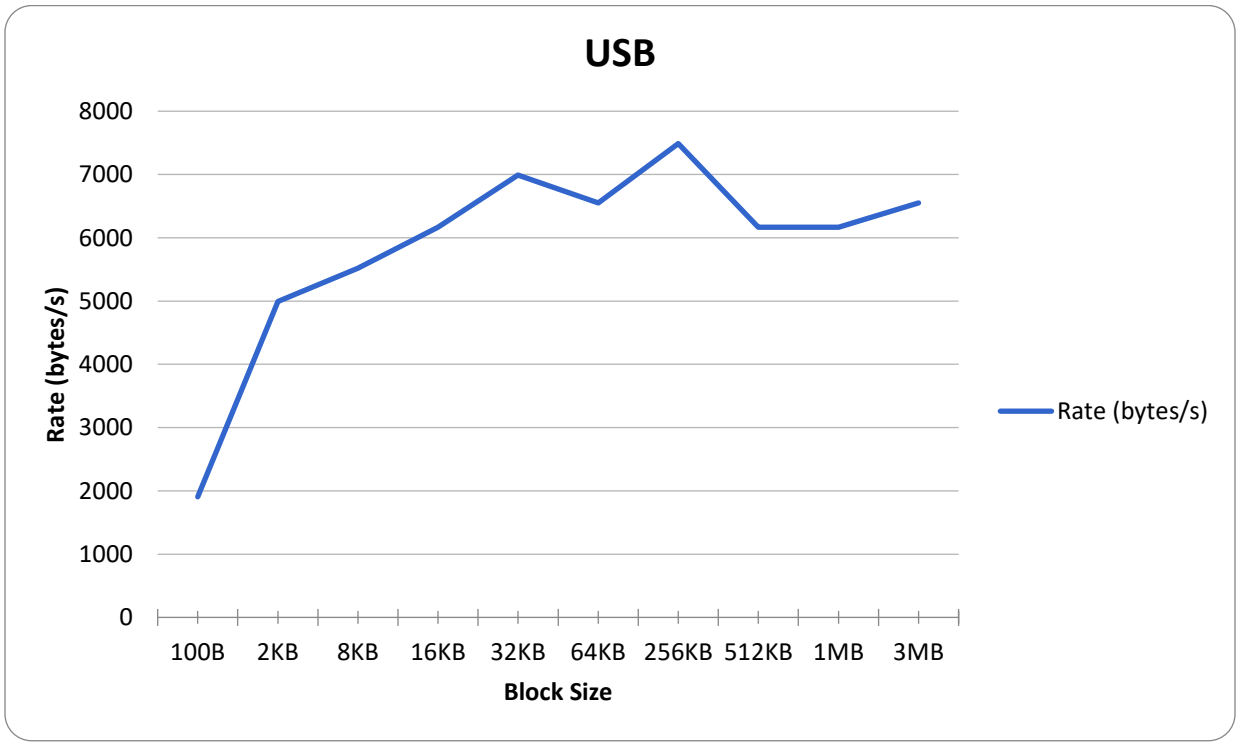
We observed that both storage mediums increased drastically in rate until block size and then the rate level off. The HDD reached the plateau-point at 16kb block size for the HDD and 8kb block size for the USB. The HDD's plateau point was about 2500 bytes/second, a lot faster than the USB which plateaued at around 1500 bytes/second.

The smaller block sizes were very slow. After the plateau points, the rates do not change significantly for different block sizes. The reason for this is because the smaller buffer sizes require more I/O cycles in order to read the entire file. On the other hand, once a certain block size was reached, the rate does not change much because the cost difference between loading a very large block size into memory and loading more of slightly smaller block sizes is likely very similar.

The optimal block sizes should be larger than 16kb. This block size loads allows for enough data to be loaded into memory without causing too much I/O. We did not encounter a large enough block size that caused the rate to drop significantly.

get_histogram





Observations:

We observed that both storage mediums increased drastically in rate until block size and then the rate level off. The HDD reached the plateau-point at 16kb block size for the HDD and 32kb block size for the USB. The HDD's peak rate was about 9500 bytes/second and then drops to a plateau at around 8000 bytes/second. For the USB, the peak rate was at around 7500 bytes/second and then drops to a plateau of around 6500 bytes/second.

The smaller block sizes were very slow. After reaching the peak rates, the rates drop slightly for larger block sizes. The reason for this is because the smaller buffer sizes require more I/O cycles in order to read the entire file. On the other hand, once the peak rate was reached, the rate drops because the cost of reading a large memory block outweighs the benefit of less I/O.

The optimal block sizes should be larger than 16kb. This block size loads allows for enough data to be loaded into memory without causing too much I/O.

Unlike the write case, larger block sizes caused the rates to drop after passing an optimal block size. We also noticed that the actual rates in the read case are significantly higher than in writing, as to be expected because the actual I/O of reading is less expensive than writing. Writing involves finding the appropriate blocks of memory to write to.