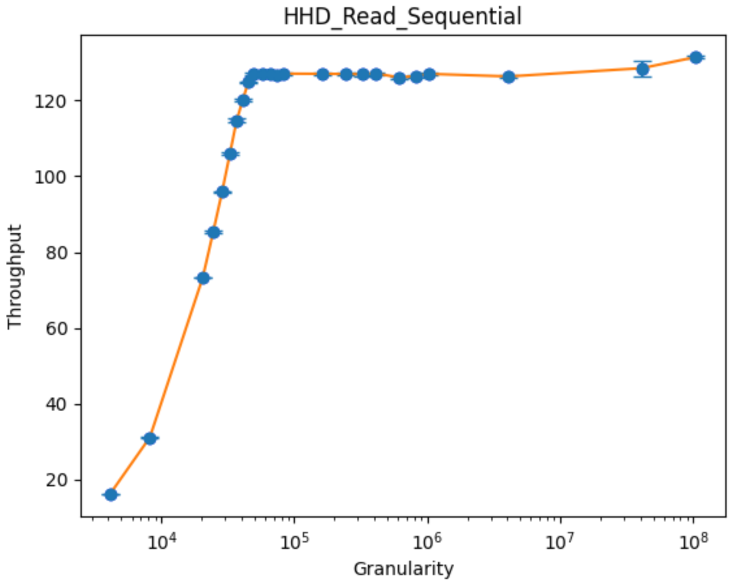
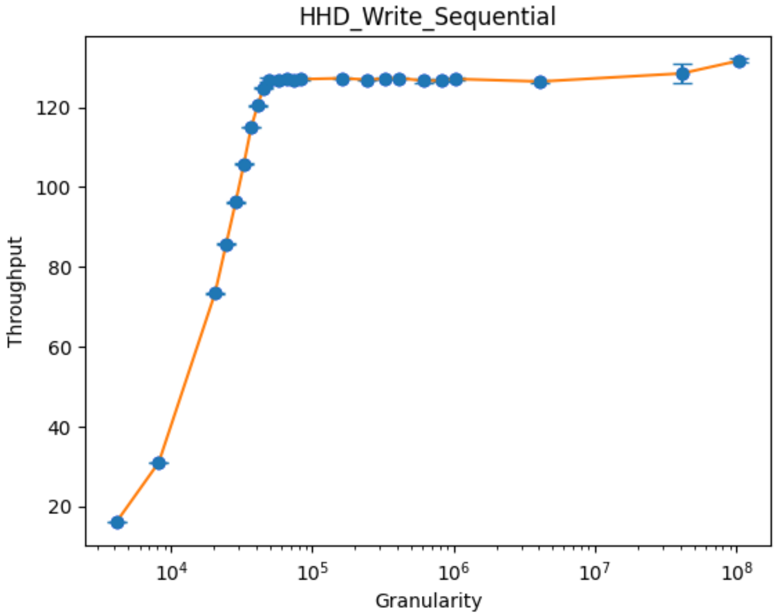
Important Disclaimer:

The SSD sda2 is gwion is bugged. When examining using ldisk, it told me that sda2 only had 1kb of space, which explains why I am getting not enough space error.

HDD_Sequantial

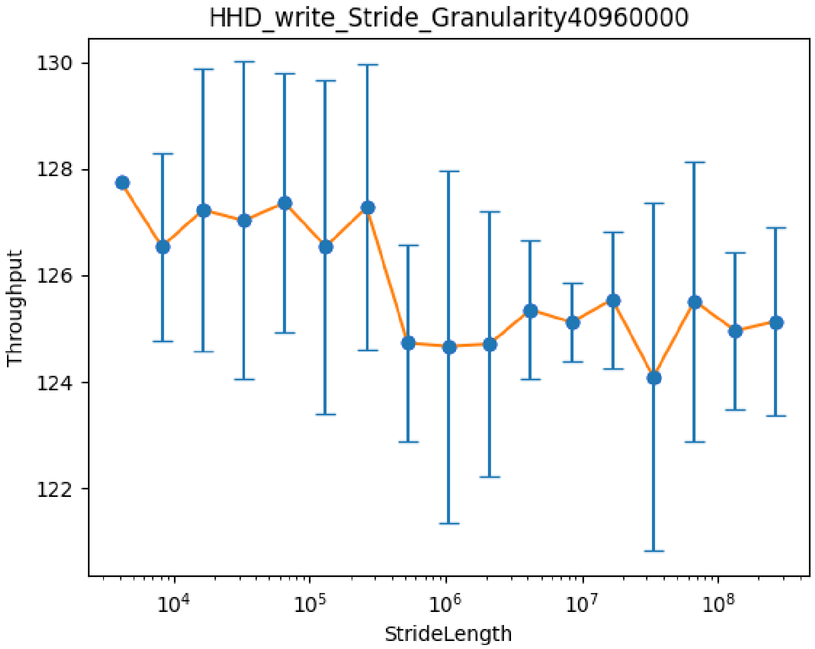


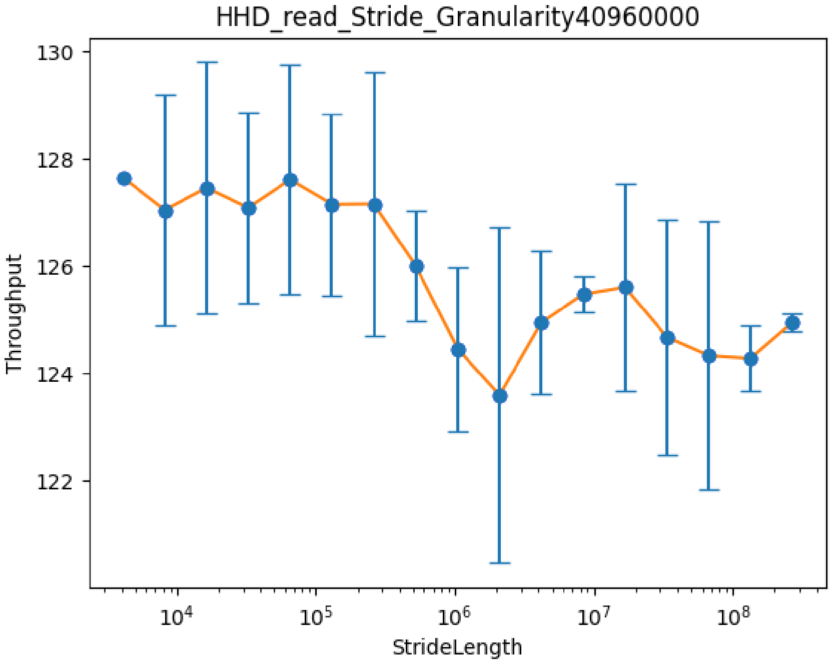


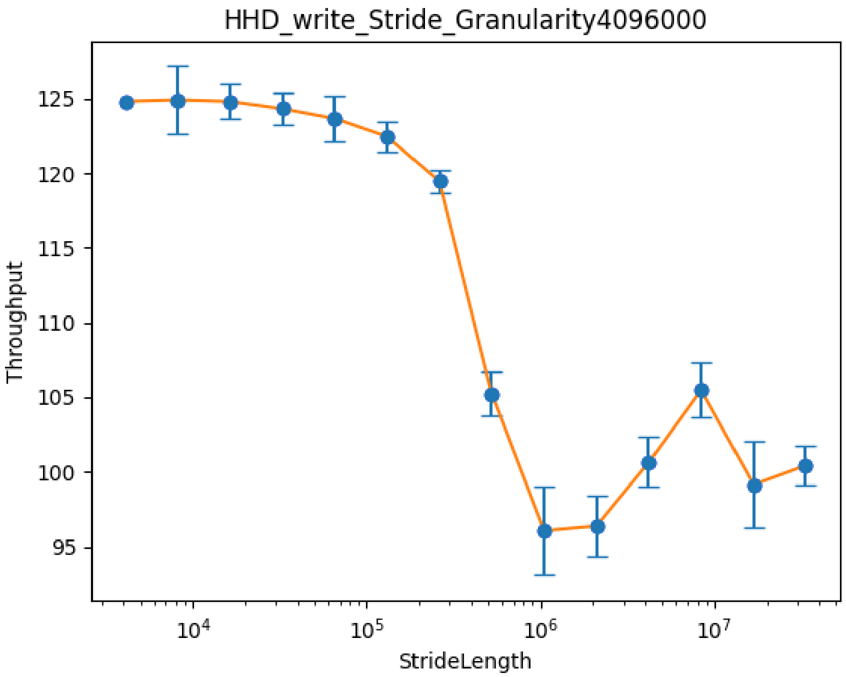
We know that throughput for HHD is, approximately, determined by disk seek time, latency, and transfer time. In sequential writing, seek time is roughly non-existent, so throughput is determined by latency and transfer time. Latency and transfer time are determined by the number of I/O operations, which is why we witnessed that rising trend at start - as we enlarge the granularity, the number of needed I/O operations to reach 1GB decreases.

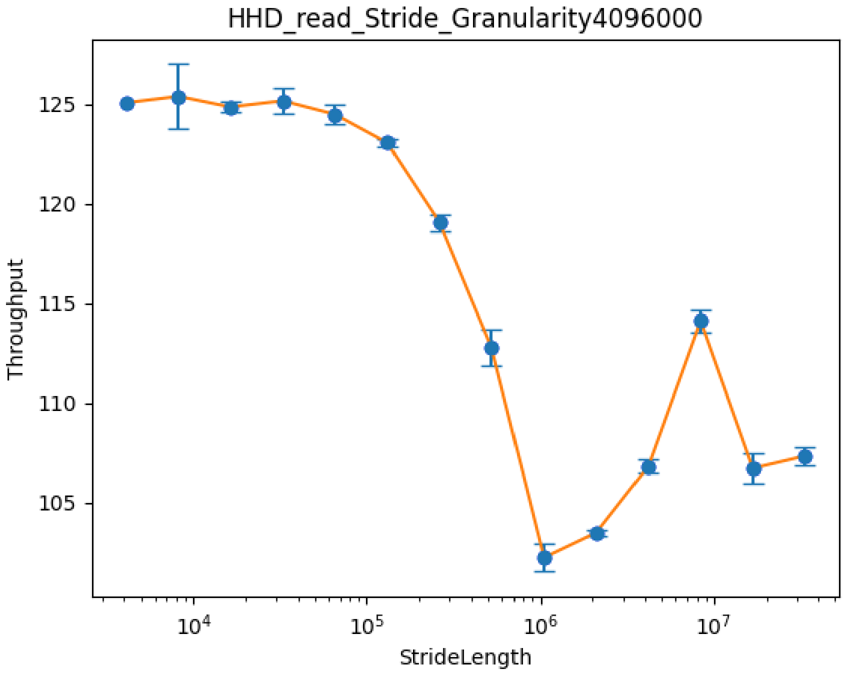
But why do we have a plateau after reaching a granularity of 12*4096? My hypothesis is that 12*4096 is the "bandwidth of our disk." I/O with granularity larger than 12*4096 will need to be completed in multiple disk writes/reads of granularity 12*4096 anyways. I guess that some optimization, such as padding with data from the "next" write/reading operation, is involved to ensure that 12*4096 bandwidth is filled in each operation. My guess is based on the fact that the throughput is very stable in the plateau. If no optimization is involved to guarantee that the bandwidth is fully utilized, we would not see such a stable line.

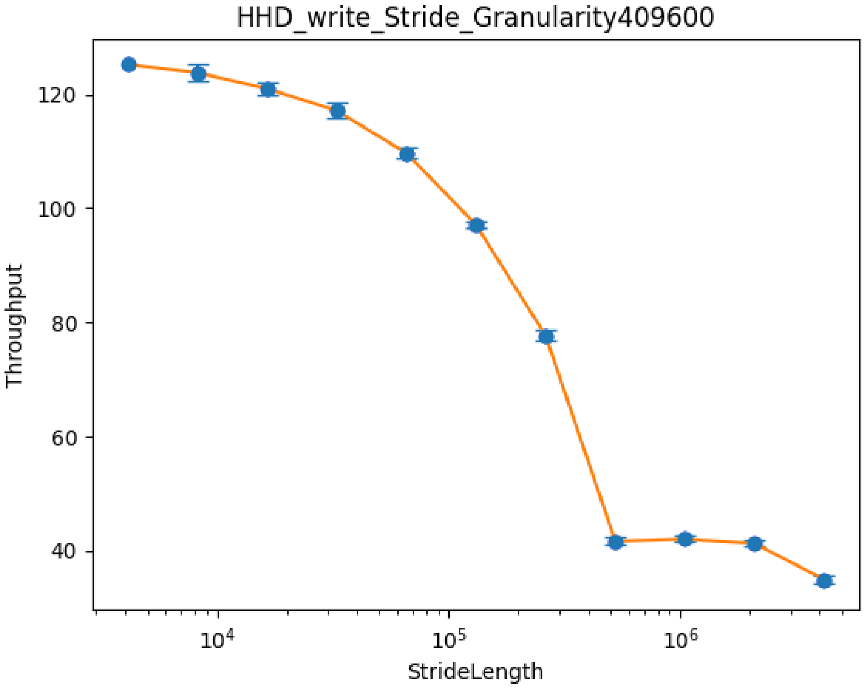
Stride Mode

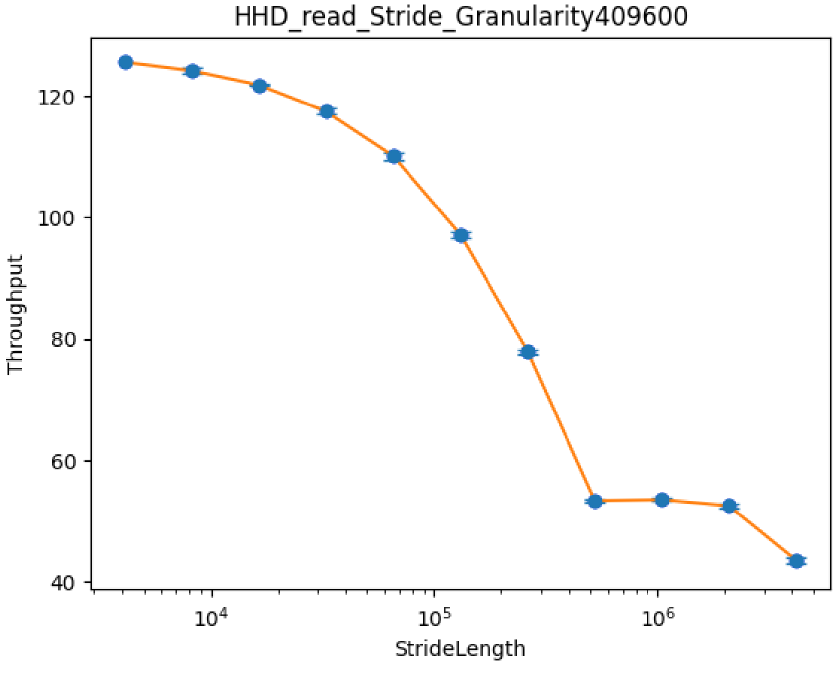

HHD_write_Stride_Granularity40960000
Throughput
130
128
126
124
122
10^4
10^5
10^6
10^7
10^8
StrideLength


HHD_read_Stride_Granularity40960000
Throughput
130
128
126
124
122
10^4
10^5
10^6
10^7
10^8
StrideLength


HHD_write_Stride_Granularity4096000
Throughput
125
120
115
110
105
100
95
10^4
10^5
10^6
10^7
StrideLength

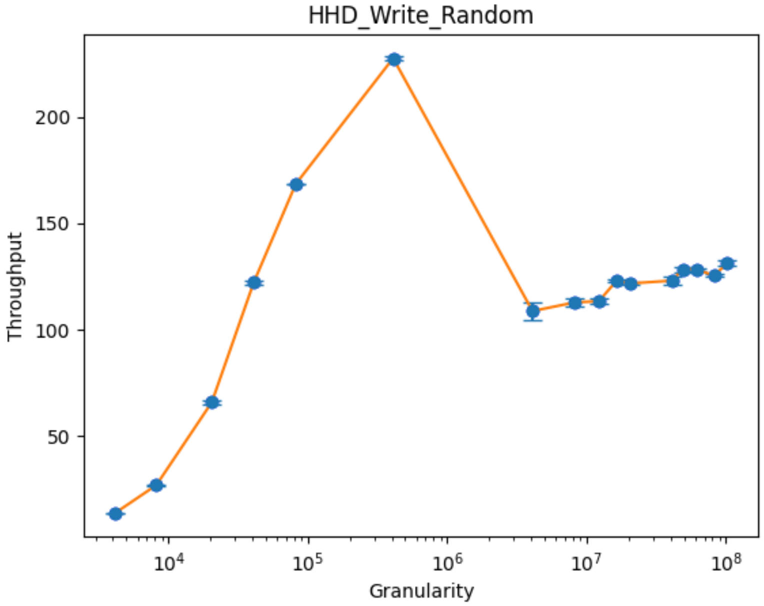

HHD_read_Stride_Granularity4096000
Throughput
125
120
115
110
105
10^4
10^5
10^6
10^7
StrideLength

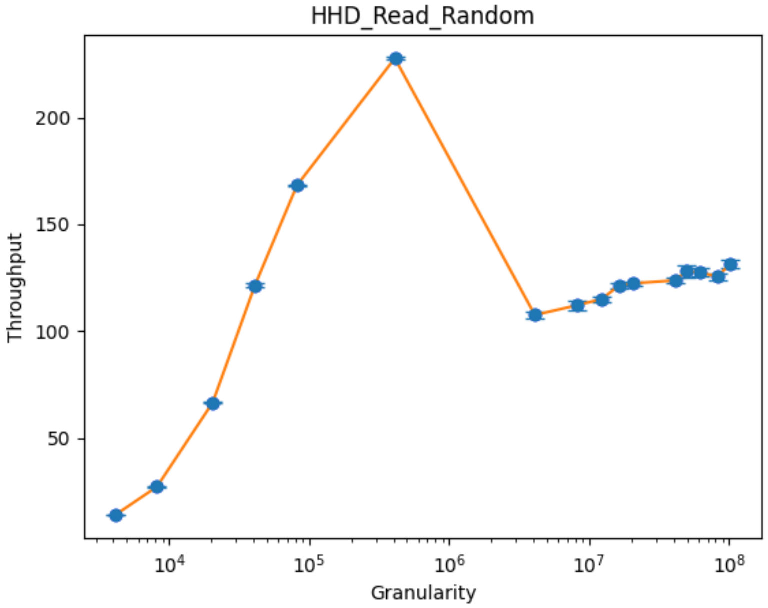

HHD_write_Stride_Granularity409600
Throughput
120
100
80
60
40
10^4
10^5
10^6
StrideLength


HHD_read_Stride_Granularity409600
Throughput
120
100
80
60
40
10^4
10^5
10^6
StrideLength

From the above graphs, it is safe to say that a high stride length will lead to lower throughput. I am not sure about why when stride length is roughly 10^7, there is a sudden increase in throughput for read_granularity4096000. My guess is that it is an outlier.

HDD_Random





I am a little surprised at the spike at middle granularity. I am guessing this has to do with the disk scheduling algorithm that gwion's system is using, but since I don't know what algorithm it is, I cannot comment on why this behavior is occurring.

I was expecting to see something similar to sequential writing mode with more variance.

SSD

I could not experiment on Gwion's ssd because it somehow has only 1kb of space, as I mentioned above. But in general, SSD has a very low seek time, so my guess is that random and stride modes would have a graph quite similar to sequential mode. Sequential mode's plot should look similar to that of HDD (an increasing trend at start due to better bandwidth utilization).

General Discussion

I don't think there is an ideal granularity, but from the experiments we can deduce a reasonable range of granularities to choose from. For instance, a granularity of less than 12*4096 is bad because it does not use all of the disk bandwidth. A larger granularity generally leads to performance increase as it not only lessens CPU time needed to process I/O, but also reduces the number of I/O operations, but as we discussed in class, it takes effort and time to optimize and fill up the large granularity. Padding to fill up granularity could lead to severe performance degradation. If I were to design a file system with a lot of I/Os, I would try to group them into granularity of above 12*4096, but not too large.

From the random access mode experiment, it seems that perhaps due to the file system disk scheduling algorithm, a performance spike could occur at "middle" granularity. I probably need to ask you about why it is behaving like that.

From my experiments, it seems that reading and writing are the same. I could not do experiments on SSD, but I suppose that SSD would be much faster.