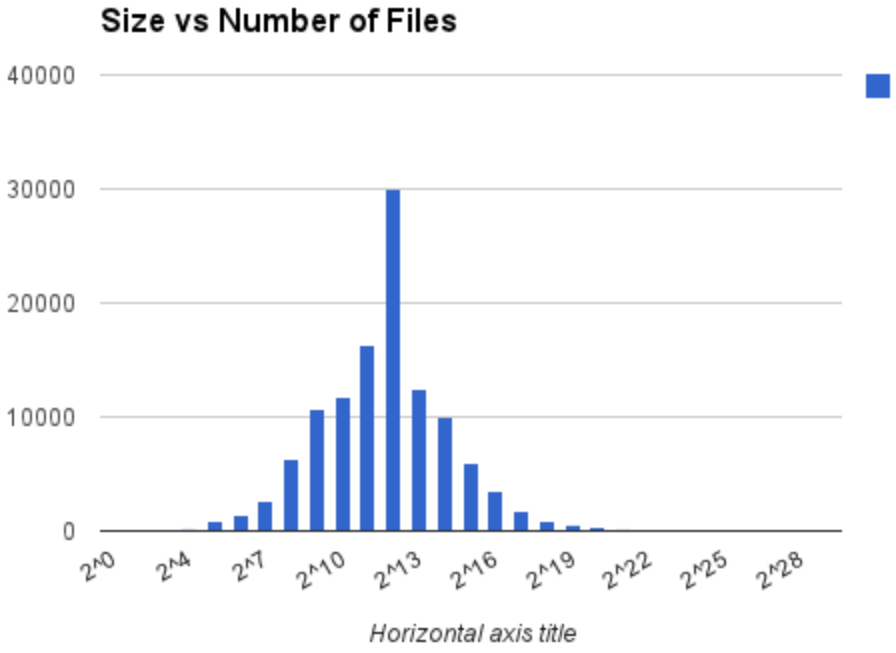
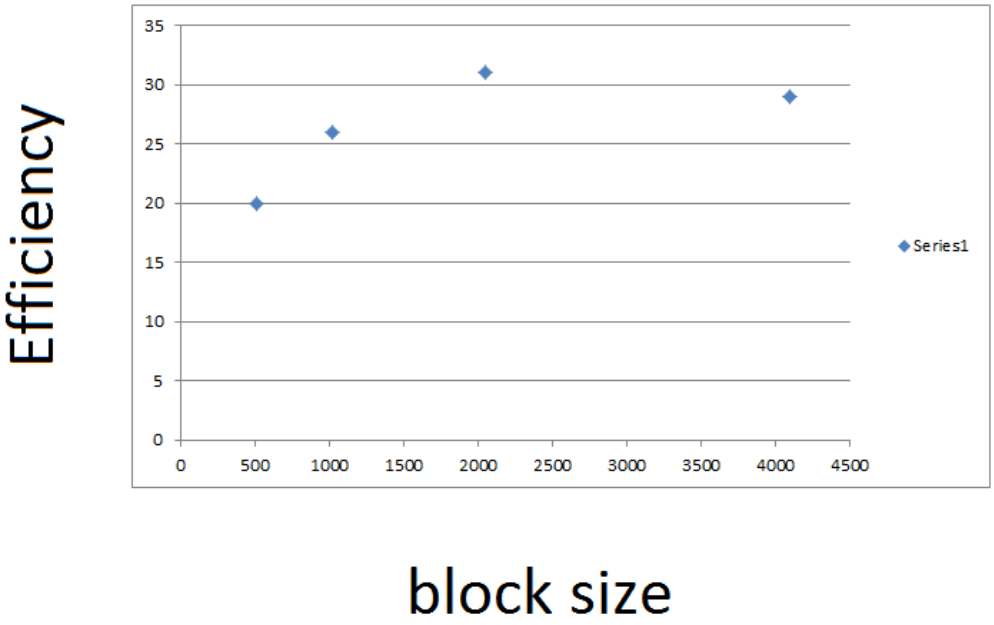
Ryan Sprowles

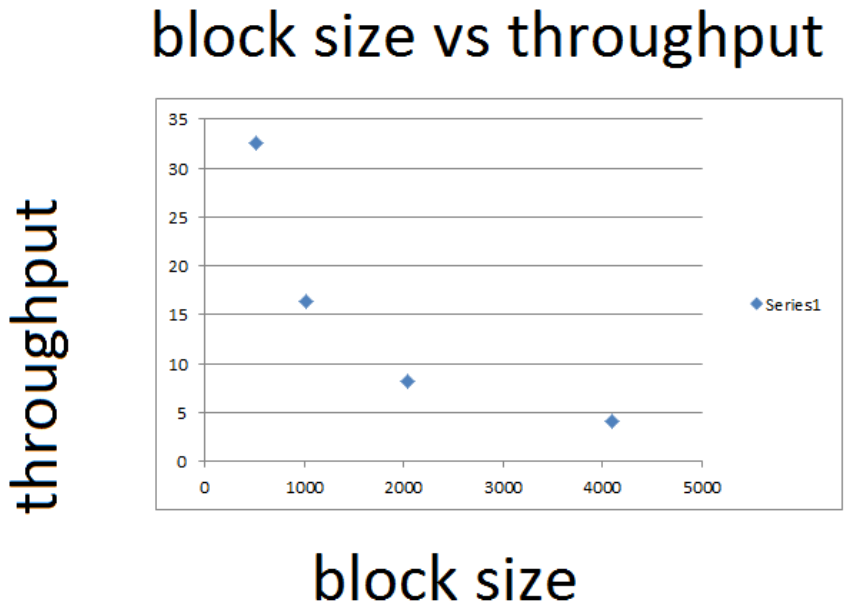


This chart give the number of files versus the file size in number of bytes.
I got this chart by running on linux the command to return all file sizes (i chose to exclude 0) and then I binned the sizes based on base 2. Most files seem to fall between 2^13 and 2 ^14, about 4 to 8 kilobytes. I did not understand what the second part was asking and was too timid to ask.



Ryan Sprowles

I apologize for the fact that its a scatter plot, i was having trouble with getting an actual graph. I did the percentages, and 2048 seemed to be the best block size in terms of efficiency for disk-space waste.



For this I used a hard drive with size 500 GB, Seek time 2.19 milliseconds, Rotational Latency 4.16 milliseconds, and Transfer Rate of 600 MB/sec. I ran this on multiple files and this was the average of the outputs. I ran with file sizes <512, 512-1023, 1024-2047,2048-4095, 4096+.

Going by the throughput and the Wasted space, the optimum block size is either 2048, or 4096. 2k had better wasted space efficiency, but the throughput for 4k was faster. I'd say that 2k would be optimum because 1 millisecond isn't much time at all, and it makes up for that with its wasted space optimization in my opinon