Assignment 3 - File System Performance

Your goal in this assignment is not to implement any major new code, but to investigate the performance of a computer's file system. This is intended to help give perspective on I/O speed before we start tackling that material in the main course. The question we are attempting to answer could be described as: "How quickly can a file system handle requests from a user process?" This is rather vague, so specifically, you are expected to offer an answer for one question:

"What is the impact of user workload on file system performance?"

In this question "user workload" refers to the requests being made upon the filesystem through the system call interface. The question asks whether differences in such a request pattern result in differences in the filesystem performance, and to what degree. This requires that you start by deciding on a characteristic that would mark two workloads as being different. A simple example of this is the degree of randomness or sequentiality of the workload. In that case, one way to answer this question is to generate files of different sizes, and to read those files in their entirety using a sequential pattern (reading a file from beginning to end), and then repeating the experiment for a random pattern (reading a bytes from random locations by making use of the "lseek" call). You could generate such test files by writing random data to a set of files, either through a very simple C program or through use of Unix command line tools (like piping the output of /dev/random through the "head" command, or by using the "dd" command instead). You are free to use whatever tools and code you wish to answer this question, but all submitted materials should be testable on lab systems and included as part of your submission where possible.

Good Luck!