LAB 8

NIKHIL KUMAR L - 140050037
UDAY KUSUPATI - 140050048

DESIGN:

We will store two files named blockstore and hashstore in the /tmp folder. The file blockstore contains the data blocks contents. The file hashstore stores the hash values. The ith line in the hashstore file corresponds to the hash value of the ith block in blockstore.

IMPLEMENTATION:

So to write a file, we first divide the file into 4KB blocks, then compute the hash value corresponding to the block, then check if it is in the hashstore file. If yes, then I just write the block number(the line number in the hash file) of that block into the file. If no, then I put the hash into the hashstore and the data block into the blockstore, and also write the block number into the file content. I write the block number in 8 bit format followed by an end line. The number 8 comes from the constraints on file size and number of files mentioned in the lab problem statement.

To read a file, I look into the block numbers which are stored in the file, then fetch that corresponding block from the blockstore file.

TEST CASES:

I have created a file a.txt with some content, and another file b.txt with the same content as a.txt followed by some other content. I.e they have 1 block in common. So when you write the second file b.txt after writing a.txt, only the second block of data in b.txt should be stored in the blockstore file, and the first block's (which is common with the a.txt data block) block number is stored from before.