ECE 469 Spring 2017 Laboratory #5: File Systems

Due Wed May 5th, 2017 at 23:59 EST, No Extension!

Introduction

The purpose of this lab is to become familiar with file systems and their implementation. In particular, you will accomplish the following in implementing a UNIX-like file system called DlxFS (DFS for short):

- implement a DFS system driver,
- implement 'newfs' to format a disk with the DFS filesystem,
- implement common libc-style file I/O functions,
- and implement multi-level directory support.

The key to successfully finishing this lab is to follow all the keys that have been given in previous labs: checking return values, well-thought-out testing plans, clear debugging statements, and lack of procrastination. In order to complete grading of this lab prior to the semester's end, we will be relying mostly on automated grading scripts, so the opportunity for partial credit is limited. These scripts require you to maintain the original directory hierarchy. Moreover, you should modify only the files mentioned in each problem.

Time Management: Please start early, as you have learned from doing Lab 4, you should not expect to finish lab 5 in the last week.

Important: Please note in PART I of this lab, without multi-level directory support, we assume the filename of a file is stored in its inode. This is an unreal simplification for this lab only. In reality, filename is not stored in its own inode. It is stored in the datablock of its parent dir file.

Background and Review

File Systems in General

A file system is simply a means of organizing a large set of bytes using a small amount of overhead, both in terms of size and access speed. There is no restriction as to what type of hardware a file system runs on: it could be Read-Only memory like a DVD-ROM, it could be a traditional magnetic or solid-state hard drive, it could be a RAM-based file system that is entirely in volatile memory, or it could even be a virtual disk that resides in a file (as what you will be implementing in this lab). In this way, the file system can shield the low-level details of hardware disk manipulation from higher-level programs.

File System Terms: Paging Revisited

A file system is similar in nature to memory paging. There is a concept of a *file*, (similar to the concept of a process's *virtual address space* in paging) which is a set of contiguous bytes. However, as you learned with paging and segmentation, it is difficult to store sets of contiguous bytes directly. Therefore, the filesystem breaks both the files and the disk up into *blocks* (analogous to virtual pages and physical pages), and keeps a table (similar to a page table) that translates the virtual block number of a file to a block on disk (or a block in the file system, this subtlety in terms will become clear soon).

Recall that with paging, you needed to store the page table and other information about a process's virtual address and physical page usage. You stored this information in a *page table*, and the pointer to the page table is stored *process control block* (PCB) structure. For file systems, we need a similar type of structure to store the information about a file: things such as the filename, the block translation table, length of the file, etc. This structure is called an *index node* or *inode* for short. An inode is the metadata associated with a file, meaning that each file has one inode, and each inode represents one file.

In lab 4, you kept track of which pages were free and which were in use through the use of a bitmap where each bit corresponded to 1 page. This was called the *freemap*. We will use the same concept in this lab to keep track of which file system blocks are in use and which are free, except that we will call it the *free block vector* in this lab instead of the freemap.

The table below summarizes the relationship between paging terms and file system terms:

File System Term	Paging Equivalent
block	page
file	process virtual address space
inode	page table/process control block
free block vector	freemap
block number	page number
physical block	physical page
virtual block	virtual page
file system block	<no equivalent>

One major difference between paging and file systems is that with paging, we had the luxury that when the system turned off, we had no need to save any information about the state of memory. It was assumed that when the machine turned on, it would rebuild all the paging infrastructure from scratch again. With file systems, this is not the case. Therefore, we must write all of this file system information to the disk before exit, and read back in when the operating system starts.

One of the primary issues when dealing with a physical disk is access time. It can take many thousands of times longer to read a byte from a physical disk than it does to read a byte from memory. For this reason, file systems generally employ some level of caching of disk data in memory. For instance, since inodes may be accessed frequently, it makes sense to keep them in memory to allow for fast access.

As with paging, there are several important terms involved with file systems whose definitions must be clear:

- **physical block**: a consecutive chunk of bytes on a physical disk. It is identified by the physical block number. This is the native number of bytes that can be read or written at a time to the physical disk.

Note that in practice this is often called a *logical disk block*, as modern disks (e.g. SCSI) can mask off bad blocks, and present to an operating system the abstraction of a one-dimensional array of consecutively numbered logical blocks.)

- **file system block**: this differs from a physical block in that a filesystem can be formatted to use blocks that are a different size than the physical disk. Its read/write functions must account for the difference in size. For instance, if a file system is formatted to use 1024-byte blocks and the disk uses 512-byte blocks, then block number 5 in the filesystem will actually correspond to blocks 10 and 11 on the physical disk.
- **virtual block**: a virtual block is used by inode-based functions. Similar to virtual addresses in paging, a virtual block number is translated to a file system block number using the translation table stored in the inode.
- **inode**: a file-system-specific structure that stores information about files.
- **free block vector**: a bitmap representing which file system blocks are in use.
- **superblock**: a special structure that stores information about the formatting of a given file system on a disk. This includes things such as where the inodes are stored, where the free block vector is stored, how many inodes there are, etc.

One important point to note is that the physical disk is only able to read or write exactly one physical block at a time. This means that if a program wants to overwrite 4 bytes in a file, and the physical block size is 512 bytes, the file system first has to read the existing 512 bytes from the disk, change the 4 bytes in the copy of the block in memory, then write the entire 512 bytes back to the disk.

Disk functions vs. DFS functions vs. DfsInode functions vs. File functions

It is easy in this lab to get confused between disk operations, file system operations, and file operations. Disk operations are performed by the actual physical disk hardware. These operations include reading and writing blocks, and reporting disk information such as the total size and the block size.

File system operations (a.k.a. DFS operations) are performed by the file system driver in the operating system. DFS operations include opening the file system (i.e. loading it into memory), closing the file system (i.e. writing it back to the disk), allocating and freeing DFS blocks, reading and writing DFS blocks, reading and writing virtual blocks (through inodes), and block caching.

File operations are performed by a special file library outside the file system driver. File operations include fopen, fclose, fread, fwrite, and fseek. These functions maintain a concept of a "current position" in a file, as well as provide an abstraction for programs that allows them to perform file I/O without having to know anything about inodes.

DLXOS File System (DFS)

The DLXOS File System uses a configurable block size, with the restriction that the file system block size must be an integer multiple of the physical block size. This keeps all file system blocks aligned with physical blocks. The file system layout is configurable based on the total physical disk size and the maximum number of inodes. An example structure for a 16MB physical disk, physical block size 512 bytes, file system block size 1024 bytes, inode size 128 bytes, and a maximum number of 128 inodes, is as follows (all block numbers are file system block numbers):

- Block 0: master boot record and superblock. Refer to the paragraph below for details.
- Blocks 1 to 16, inclusive: array of inode structures.
- Blocks 17 to 18, inclusive: free block vector

- Blocks 19 to 16383, inclusive: data blocks

File system block 0 is actually represented as two sections: the first section is the "master boot record", where a normal file system would store information necessary to boot an operating system, and the second section is the superblock structure. The location of the master boot record is an agreement between the machine hardware/firmware (e.g. BIOS) and the operating system. Therefore, it must always reside in physical block 0 (note *physical block 0*, not file system block 0, since the hardware knows nothing about our file system). Since we won't be booting from this disk, we will write all zeros to this block.

Also, since the file system driver must read the file system blocksize from the superblock (recall this is configurable), then it must also know exactly which physical block contains the superblock. After reading the superblock, it can then operate in terms of file system blocks. Therefore, the superblock will always reside at physical block 1, right behind the master boot record. The inodes start in the first file system block (NOT physical block) after the superblock, and the free block vector starts in the first file system block after the inode blocks.

To make debugging simpler, a DFS inode structure must be exactly 96 bytes large. Since the filename for an inode is stored in the structure, you can adjust the maximum filename length such that the overall size of the inode is exactly 96 bytes.

The following items should be stored in the superblock (you may add more if you like):

- a valid indicator for the file system
- the file system block size
- the total number of file system blocks
- the starting file system block number for the array of inodes
- the number of inodes in the inodes array
- the starting file system block number for the free block vector.

The following items should be stored in the inode structure (you may add more if you like, just remember that the total size must be 96 bytes):

- an in use indicator to tell if an inode is free or in use
- the size of the file this inode represents (i.e. the maximum byte that has been written to this file)
- the filename, which is just a string
- a table of direct address translations for the first 10 virtual blocks
- a block number of a file system block on the disk which holds a table of indirect address translations for the virtual blocks beyond the first 10.

The direct address translation table in the inode contains file system block numbers for allocated file system blocks that belong to this file. Since most files in any file system tend to be very small, keeping the first 10 translations in this table makes the address translation faster for most files.

In order to keep the size of the inode small, however, large files are going to need additional space beyond the first 10 virtual blocks. Therefore, each file that grows beyond the first 10 virtual blocks will allocate an entire file system disk block to store the rest of its translation table. In DFS, a file is not allowed to grow larger than is supported by the maximum number of entries in both the direct and indirect tables.

The DFS driver is in charge of which information is loaded into memory and which is read from the disk on demand. For simplicity, we do not deal with filesystem reliability in the presence of crash failure in this lab. Specifically, to support fast access, when the operating system starts, the DFS driver should load the superblock, all inodes, and the free block vector into structures in memory. It should then mark the superblock on the disk as invalid. When the operating system exits, it should write all the inodes and the free block vector back to the disk, and then write the superblock back to the disk marked as valid. If the system crashes before writing the file system information back to the disk, then the file system is considered corrupted.

Multi-level Directory Support

So far, we have been discussing a flat directory structure where all files are essentially in the same directory. In the last task of this lab, you will add multi-level directory support and userland interfaces to the file system in DLXOS.

Directories

At the system level, directories and files should be treated almost identically. The same handlers written above should be used for both files and directories. Directories distinguish themselves from files in that they always have a predefined structure. The directory structure for this lab will be a sequence of fixed-length directory entries. Each directory entry should be 76 bytes wide. The first 72 bytes will contain the filename and the final four bytes will contain the inode number corresponding to the file. This means that a filename can be at most 72 bytes long. If the filename is less than 72 bytes, it should be null terminated (using `'\0'`).

Additionally, you should use `'\0'` to mark free directory entries. Any entry having `'\0'` as its first character should be considered empty and therefore available. Thus in order to remove a file (or a directory) from the file system, you should set the corresponding directory entry to `'\0'` and release any associated resources. To create a file or directory, simply create an inode with the proper attributes and add an appropriate directory entry.

You may assume that the depth of the directory structure will not exceed 10. You don't have to assume this, but we will not test your code beyond a depth of 10. This may or may not simplify your life. Finally, The `'/'` character should not be permitted as part of a filename, for obvious reasons.

Permissions

For this project our file permissions are similar to those used in UNIX with the exception that we do not have group permissions (ie, we only have owner and other). You can assign permissions when you create a file or directory. File permissions should be stored as an 8-bit vector as follows:

```
Bit:        5  4  3  2  1  0
Permission: UR UW UX OR OW OX
```

Where `U` is for the user, or owner of the file and `O` is for other (any process *except* the owner). Permissions for a given directory or file are stored in its corresponding inode using the format mentioned above. Following the UNIX convention, in order to change over to a directory you must have execute permission for that directory. To open a file, the calling process therefore at least requires execute permission to all of the directories leading to the file in addition to proper permissions on the file itself.

Thus, to open the file `/a/b` for reading, the process must have execute permission on `/` and `a` and read permission on `b`. To change over to the directory `/a/c`, the process requires execute permission on all three directories - `/`, `a`, and `c`.

File deletion and creation (as well as directory deletion and creation) requires write access to the parent directory. Thus, to delete a file or directory you must have both write and execute permission to its parent directory along with execute permission on all directories leading to the parent.

As already mentioned, permissions are divided into two sets - user (`u`) and others (`o`). To check permissions, you should assume that the user id is the running process' pid for simplicity. You can compare the pid with a file's (or directory's) ownerid. The ownerid should always correspond to the pid of the process that created that file.

Changes to DLXOS Source

Physical Disk Simulator

We have written a physical disk "simulator" for you in os/disk.c and include/os/disk.h. This simulator uses built-in functions in filesys.c that read and write to an actual file on the Linux system as a disk image. It supports the following functions:

- int DiskBytesPerBlock(): this function returns the number of bytes in one physical block.
- int DiskSize(): this function returns the number of bytes in the entire disk.
- int DiskCreate(): this function creates the Linux file that will hold the file system.
- int DiskWriteBlock(uint32 blocknum, disk_block *b): this function writes one physical block of data to the file at the specified block number.
- int DiskReadBlock (uint32 blocknum, disk_block *b): this function reads one physical block of data from the file at the specified block number.

IMPORTANT: the name of this Linux file is #define-d in include/os/disk.h. You must change this name before the simulator will work. Your file system file must reside in the /tmp directory, and must be named ee469gXX.img, where "XX" is your group number. The reason this is required is that the /tmp directory resides on the local hard disk of the computer you are logged in to. Your home directory resides on an NFS file share on a remote server. If all your groups use a file in your home directory as your file system, the NFS server has been known to be slowed down to a halt (e.g. as if under a DDoS attack) as you all keep writing and reading the large "disk" files over and over, especially the night before the assignment is due. Therefore, you have to store the file in a local location. Since you are storing it in a local location, if you all use the same file name, then if two groups are logged in to the same physical machine, the two groups will keep overwriting each other's file (or keep getting access permission denied). Therefore, you have to change the filename to be in /tmp and to be unique to your group.

Blockprint Script

Your primary means of debugging in this lab will be to look at the bytes in your filesystem file to see if they are correct. To do this, you must have the ability to easily look at a binary file. To that end, we have written a simple bash script that uses the xxd utility and sed to allow you to look at specific "blocks" of a binary file, specified by block number. This script resides in the scripts/ directory. You do not have to use this script: it is only provided for your convenience. You can use the script in one of three ways:
$ blockprint /tmp/ee469g98.img
This will print all the blocks of the file /tmp/ee469g98.img to the screen.
$ blockprint 1 /tmp/ee469g98.img
This will print physical block 1 of /tmp/ee469g98.img to the screen (the superblock).

$ blockprint 34 37 /tmp/ee469g98.img
This will print physical blocks 34 to 37, inclusive, to the screen (the free block vector).

Note that the block size is configurable by changing a variable in the script if you prefer to print DFS-sized blocks instead of physical-sized blocks.

Operating System Tests

We added a convenient user trap, application, and os source file for enabling a user program to run a set of tests within the operating system. Since there is a lot of work to be done in this lab before you can use a user program to read/write files, we require you to test some intermediate functions (such as inode-based functions), which are not visible to user programs. Therefore, the apps/ostests user application is a one-line program that calls the run_os_tests() trap. This triggers the function in os/ostests.c called "RunOSTests()". You can put any sequence of internal OS testing code here that you like, then run the ostests user program to see it run.

GracefulExit replaces exitsim()

We have replaced all the calls to "exitsim" in the simulator code with calls to "GracefulExit", which is a new function defined in traps.c. This function simply tries to first close the filesystem, then calls exitsim. This functionality is required since the disk copy of the filesystem is invalidated until the file system is closed.

New OS Files for Your Code

You will need to fill in os/dfs.c, os/files.c, include/dfs_shared.h, include/files_shared.h, include/os/dfs.h, and include/os/files.h with your code for this lab. The dfs-named files deal with the DFS file system driver, and the files-named files deal with the file I/O interface to user programs (think fopen, fclose, etc.).

Shared OS and user header files

You will be writing two user programs in this lab: one to format a disk and one to test file reading and writing. These programs and the operating system will need to share some information such as the inode and superblock structures, file descriptor structures, etc. To facilitate this, user programs #include the shared versions of the header files (in include/), and os programs #include the os-specific versions of the header files (in include/os/), and these os-specific header files #include the shared header files at the top. You should not share any information with user programs that is not necessary, such as os-level function prototypes, internal OS #define-d constants, etc.

#define-d Constants vs. Superblock Information

It is very tempting to #define all the relevant information about the file system. For instance, the user program that formats the file system can #define the block size. It would be tempting to #define this in the include/dfs_shared.h header file, and use the same constant in the operating system. However, since the block size is configurable, if I use your file system driver on a filesystem that I created with a different block size, it should still work. If you use your #define-d constant, you will be using the wrong block size. Recall that the block size (and all other configurable file system options) are written in the superblock structure. Your driver therefore should use the values in the superblock structure after it is read from the disk, rather than a constant.

You may ask, however, then how are you supposed to know how large to make your dfs_block stuctures and inode and free block vector arrays at compile time, if you won't know how big any of them are until you read the superblock at run time? The answer is that you must assign a #define-d maximum value to each of these items, and then use however many items/bytes that the superblock requires at runtime.

When grading, we will take off points for using any #define-d constants where the superblock information could be used instead.

Multi-level Directory Support

For this part, you should implement your code in the `lab5/multilevel` directory given to you. You may wish to copy some of the functionality from the `lab5/flat` directory but DO NOT SIMPLY COPY IT! Merge the changes manually, as the functions for this part will require modifications to most of the functions implemented for the earlier parts.

Assignment

Download and untar ~ee469/labs_2017/Labs/lab5.tar.gz. This will create the `lab5/flat` and `lab5/multilevel` directories for you. Put all your work for problmes 1 - 6 in the `lab5/flat` directory structure and put your code for problem 7 in the `lab5/multilevel` directory structure. **DON'T FORGET TO CHANGE THE GROUP FILENAME FOR YOUR FILESYSTEM in `lab5/flat/include/os/disk.h` and `lab5/multilevel/include/os/disk.h`. To keep things simple, a particular file can be opened (using FileOpen()) by only one process.**
You should use locks when modifying shared data structures. Use locks in the following cases:

- whenever you allocate inodes and file descriptors
- whenever you allocate or deallocate file system blocks using the free block vector

1. (10 points) **Implement a user program named "newfs" to create a new file system**. Put your code for this in `lab5/flat/apps/newfs or lab5/flat/apps/fdisk`. You can use some existing user traps for the disk simulator that will allow you to write blocks, read the disk size, and read the disk blocksize. Your file system should be formatted as outlined above in this lab handout. Use the following options:
   - your inode structure should be 96 bytes
   - you should have 192 inodes in your filesystem
   - you should have 10 direct-addressed entries in the inode's virtual block translation table
   - you should have one indirect-addressed block that is only allocated as needed

   Note: For this lab, the maximum value of DFS block size is 1024 bytes and that of the disk block is 512 bytes. The maximum physical disk size is 16MB.

2. (20 points) **Implement a DFS file system driver in DLXOS**. Write the following functions in dfs.c. Note that the inode-based functions will be implemented later. Be sure that you minimize the amount of information that is shared between the os and the user programs through dfs_shared.h if that information could be better shared through the superblock.
   - void DfsModuleInit(): initialize and open the filesystem
   - int DfsOpenFileSystem(): read the superblock, inodes, and free block vector from the disk, then invalidate the disk's copy. Return DFS_FAIL on failure, and DFS_SUCCESS on success. This fails if the filesystem is already open.
   - void DfsInvalidate(): invalidate the current memory copy of the filesystem. You'll need this in

order to be able to format a disk.

- int DfsCloseFileSystem(): write the superblock, inodes, and free block vector back to the disk. Write the valid superblock last to make sure the other actions succeed. Return DFS_FAIL on failure, and DFS_SUCCESS on success. This fails if the filesystem is not open.
- int DfsAllocateBlock(): Allocate a file system block, returning its block number on success, DFS_FAIL on failure. This fails if the filesystem is not open. Remember to use locks whenever you allocate or deallocate file system blocks using the free block vector.
- int DfsFreeBlock(int blocknum): Free a file system block. Return DFS_FAIL on failure, and DFS_SUCCESS on success. This fails if the filesystem is not open.
- int DfsReadBlock(int blocknum, dfs_block *b): read one filesystem block from the disk, storing it in dfs_block b. Remember that one filesystem block can span multiple numbers of physical blocks. Return DFS_FAIL on failure, and the number of bytes read on success. This fails if the filesystem is not open.
- int DfsWriteBlock(int blocknum, dfs_block *b): write one filesystem block to the disk. Return DFS_FAIL on failure, and the number of bytes written on success. This fails if the filesystem is not open.

You are encouraged to write any additional functions that may help your implementation efforts.

3. (15 points) **Implement the following inode-based functions in the DFS driver**. Note that all of these functions fail if the file system in memory is not valid.
    - int DfsInodeFilenameExists(char *filename): look through all the inuse inodes for the given filename. If the filename is found, return the handle of the inode. If it is not found, return DFS_FAIL.
    - int DfsInodeOpen(char *filename): search the list of all inuse inodes for the specified filename. If the filename exists, return the handle of the inode. If it does not, allocate a new inode for this filename and return its handle. Return DFS_FAIL on failure. Remember to use locks whenever you allocate a new inode.
    - int DfsInodeDelete(int handle): de-allocate any data blocks used by this inode, including the indirect addressing block if necessary, then mark the inode as no longer in use. Return DFS_FAIL on failure, and DFS_SUCCESS on success.
    - int DfsInodeReadBytes(int handle, void *mem, int start_byte, int num_bytes): read num_bytes from the file represented by the inode handle, starting at virtual byte start_byte, copying the data to the address pointed to by mem. Return DFS_FAIL on failure, and the number of bytes read on success.
    - int DfsInodeWriteBytes(int handle, void *mem, int start_byte, int num_bytes): write num_bytes from the memory pointed to by mem to the file represented by the inode handle, starting at virtual byte start_byte. Note that if you are only writing part of a given file system block, you'll need to read that block from the disk first. Return DFS_FAIL on failure and the number of bytes written on success.
    - int DfsInodeFilesize(int handle): return the size of an inode's file. This is defined as the maximum virtual byte number that has been written to the inode thus far. Return DFS_FAIL on failure.
    - int DfsInodeAllocateVirtualBlock(int handle, int virtual_blocknum): allocate a new filesystem block for the given inode, storing its blocknumber at index virtual_blocknumber in the translation table. If the virtual_blocknumber resides in the indirect address space, and there is not an allocated indirect addressing table, allocate it. Return DFS_FAIL on failure, and the newly allocated file system block number on success.
    - int DfsInodeTranslateVirtualToFilesys(int handle, int virtual_blocknum): translate the virtual_blocknum to the corresponding file system block using the inode identified by handle. Return DFS_FAIL on failure.

4. (10 points) **Write test code in ostests.c to test the full file system driver**. You should test writing non-block-aligned sets of bytes, writing a file large enough to use indirect addressing, and verify that any written bytes are persistent between runs of the simulator.

5. (15 points) **Implement the following file-based API**. These functions should use the file system driver functions found in dfs.c. You will need to implement a file descriptor structure that stores relevant information about an open file such as the filename, the inode handle, the current position in the file, an end-of-file flag, and the mode. Note that a particular inode can be open in only one file descriptor. Also, the maximum number of bytes that can be read or written at any time by the file functions is 4096 bytes. All of these functions should only allow the process that opened a given file to use the open file descriptor.
   - int FileOpen(char *filename, char *mode): open the given filename with one of three possible modes: "r", "w", or "rw". Return FILE_FAIL on failure (e.g., when a process tries to open a file that is already open for another process), and the handle of a file descriptor on success. Remember to use locks whenever you allocate a new file descriptor.
   - int FileClose(int handle): close the given file descriptor handle. Return FILE_FAIL on failure, and FILE_SUCCESS on success.
   - int FileRead(int handle, void *mem, int num_bytes): read num_bytes from the open file descriptor identified by handle. Return FILE_FAIL on failure or upon reaching end of file, and the number of bytes read on success. If end of file is reached, the end-of-file flag in the file descriptor should be set.
   - int FileWrite(int handle, void *mem, int num_bytes): write num_bytes to the open file descriptor identified by handle. Return FILE_FAIL on failure, and the number of bytes written on success.
   - int FileSeek(int handle, int num_bytes, int from_where): seek num_bytes within the file descriptor identified by handle, from the location specified by from_where. There are three possible values for from_where: FILE_SEEK_CUR (seek relative to the current position), FILE_SEEK_SET (seek relative to the beginning of the file), and FILE_SEEK_END (seek relative to the end of the file). Any seek operation will clear the eof flag.
   - int FileDelete(char *filename): delete the file specified by filename. Return FILE_FAIL on failure, and FILE_SUCCESS on success.

6. (10 points) **Write a user application to test your file API**. Demonstrate that you can open a file, write bytes to it, close it, re-open it, and read the same bytes back.

7. (20 points) **Implement a multi-level directory structure in DLXOS**. You will add multi-level directory support and userland interfaces to the file system in DLXOS.
   - The file and path names should be specified in the same manner as in UNIX. Eg, `/a/b` would refer to file `b` located in directory `a`, which in turn is found in the root directory (`/`).
   - You should track the current working directory (cwd) for each process. We suggest using a `currentDIR` field in the PCB. You can use it to store the inode number for the cwd.
   - There are a number of helper functions and other useful information in **`lab5/multilevel/dfshelp.c`**. You should understand them and possibly use them after making modifications and copying where appropriate.
   - For simplicity, you are not required to support the special directory entries "`.`" and "`..`". This means that the path to a given file can only be specified explicitly (from the root directory) or starting with the current directory. For the latter to succeed, the file obviously must be a child of the current directory in the hierarchy.

   Functions to implement

   - int FileOpen(char *filename, char *mode): open the given filename with one of three possible modes: "r", "w", or "rw". The path is either specified from the current directory, or from the root directory. Return FILE_FAIL on failure, and the handle of a file descriptor on success.

1. **`r`**: all directories along the path should exist and be executable (note also that all names in the path except the leaf must be directories). The leaf must exist and be readable. The current position should be set to the beginning.
2. **`w`**: all directories along the path should exist and be executable. The leaf must exist, be a file, and be writable. Directories should not be able to be opened for write access (they can be read, however). The file size should be truncated to 0 (ie, all existing data blocks should be freed and the size updated). The position should be set to the beginning of the file. If the leaf does not exist, a file by that name should be created - note that for creation to succeed the parent directory must be writable.
3. **`rw`**: all directories along the path should exist and be executable. The leaf must exist and be readable and writable. The leaf must be a file. The position should be set to the beginning of the file. The file should be truncated to size 0 (see the previous mode).

Sample usage:

```
handle = fopen("/a/b/c", "r");
handle = fopen("d/e/f", "rw");
```

- int FileClose(int handle): close the file specified by handle. Return FILE_FAIL on failure, and FILE_SUCCESS on success.
- int FileRead(int handle, void *mem, int num_bytes): read num_bytes from the open file descriptor identified by handle. The file must have been opened in `r`, or `rw` mode for this function to succeed. The function should increment the current position in the file by the number of bytes read. Return FILE_FAIL on failure or upon reaching end of file, and the number of bytes read on success. If end of file is reached, the end-of-file flag in the file descriptor should be set.
- int FileWrite(int handle, void *mem, int num_bytes): write num_bytes to the open file descriptor identified by handle. The file must have been opened in `w`, or `rw` mode for this function to succeed. This function should increment the current position in the file by number of bytes written. Return FILE_FAIL on failure, and the number of bytes written on success.
- int FileDelete(char *path): delete the file specified by path. All directories leading to the leaf must be executable. The leaf node must be a file (not a directory). The parent directory must additionally be writable. The process invoking this function must also own the leaf (permissions of the leaf do not matter - only ownership). On success, this function should clear the directory entry corresponding to the leaf and free all associated resources. Return FILE_FAIL on failure, and FILE_SUCCESS on success.
- int MkDir(char *path, int permissions): create the specified directory. All directories leading to the leaf must be executable. The parent directory must additionally be writable and the leaf should not already exist. Return FILE_FAIL on failure, and FILE_SUCCESS on success.
- int RmDir(char *path): this function behaves identical to `FileDelete()` except on a directory. Additionally, the leaf must be empty. (eg, all directory entries must be `'\0'`.) If called on the root directory, this function should fail. If the leaf is not empty, this function fails. Return FILE_FAIL on failure, and FILE_SUCCESS on success.

Tips and Tricks

- For this part you are responsible for opening/shutting down the file system as well as setting up the root inode. You can do this in `main()` in `process.c`. The root inode should only be setup once (you might want to do this elsewhere). The root inode should be stored in inode 0.
- When a process is first created, its current directory should be set to `/`.

- There can only be one file or directory with a given name in a given directory. Directories and files may not share the same name in the same parent directory.
- The default permissions for file creation should be ur+uw+ux.
- You can remove a directory even if it is currently the current directory of a process. Think about what this means - all of your functions should check the inode for the current directory when it is referenced to make sure it is still marked as in use. An error should be returned if the directory no longer exists.

Turning in your solution

You should make sure that your `lab5/` directory contains all of the files needed to build your project source code. You should include a README file that explains:

- how to build your solution,
- anything unusual about your solution that the TA should know,
- and **a list of any external sources referenced while working on your solution**.

DO NOT INCLUDE OBJECT FILES in your submission! In other words, make sure you run "make clean" in all of your application directories, and in the os directory. Every file in the turnin directory that could be generated automatically by the DLX compiler or assembler will result in a 5-point deduction from your over all project grade for this lab.

When you are ready to submit your solution, change to the directory containing the `lab5` directory and execute the following command:

```
turnin -c ee469 -p lab5-Y lab5
```

where `Y` stands for your lab session, so it will be lab5-1 or lab5-2 or lab5-3 depending on which lab session you are in.
Wednesday, 2:30pm = 1, Friday, 11:30am = 2, Friday, 2:30pm = 3.

---

ee469@ecn.purdue.edu