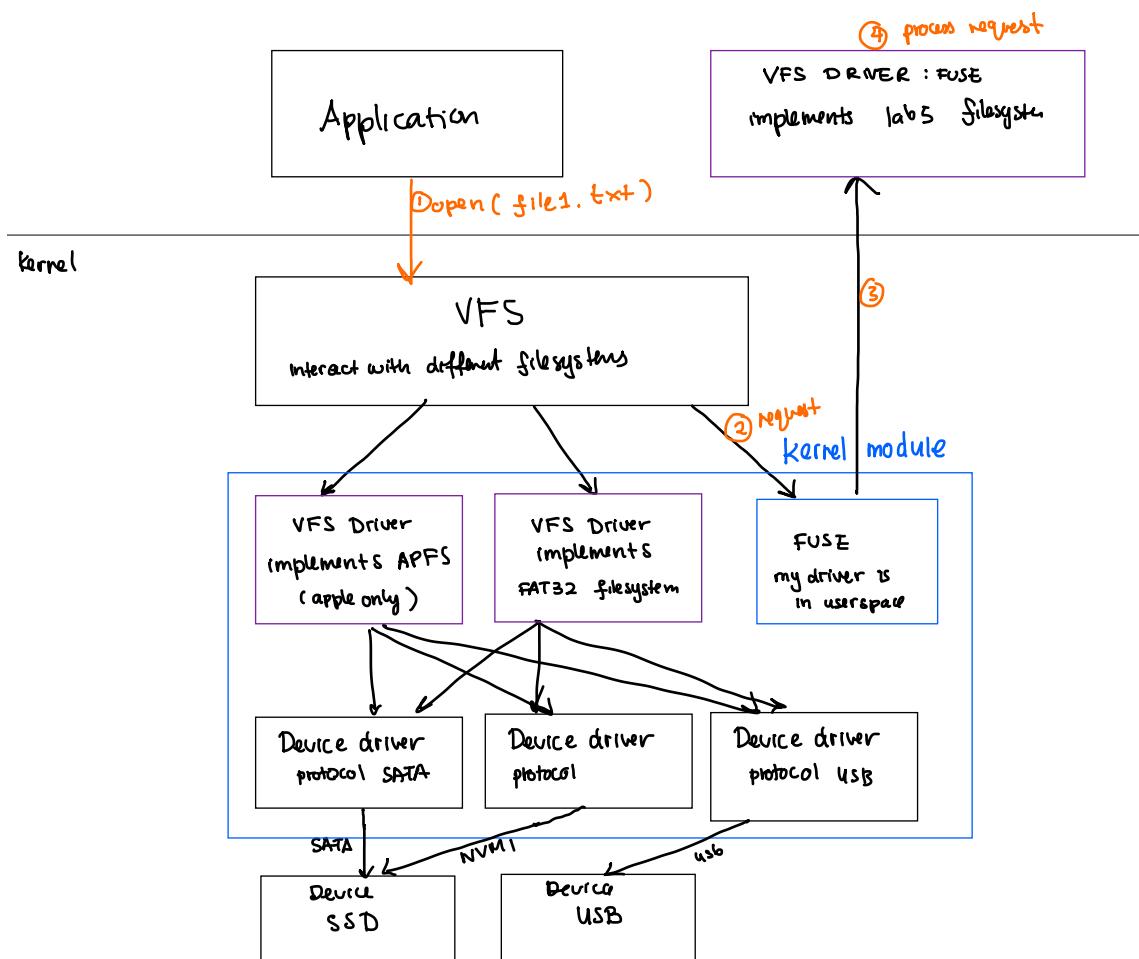


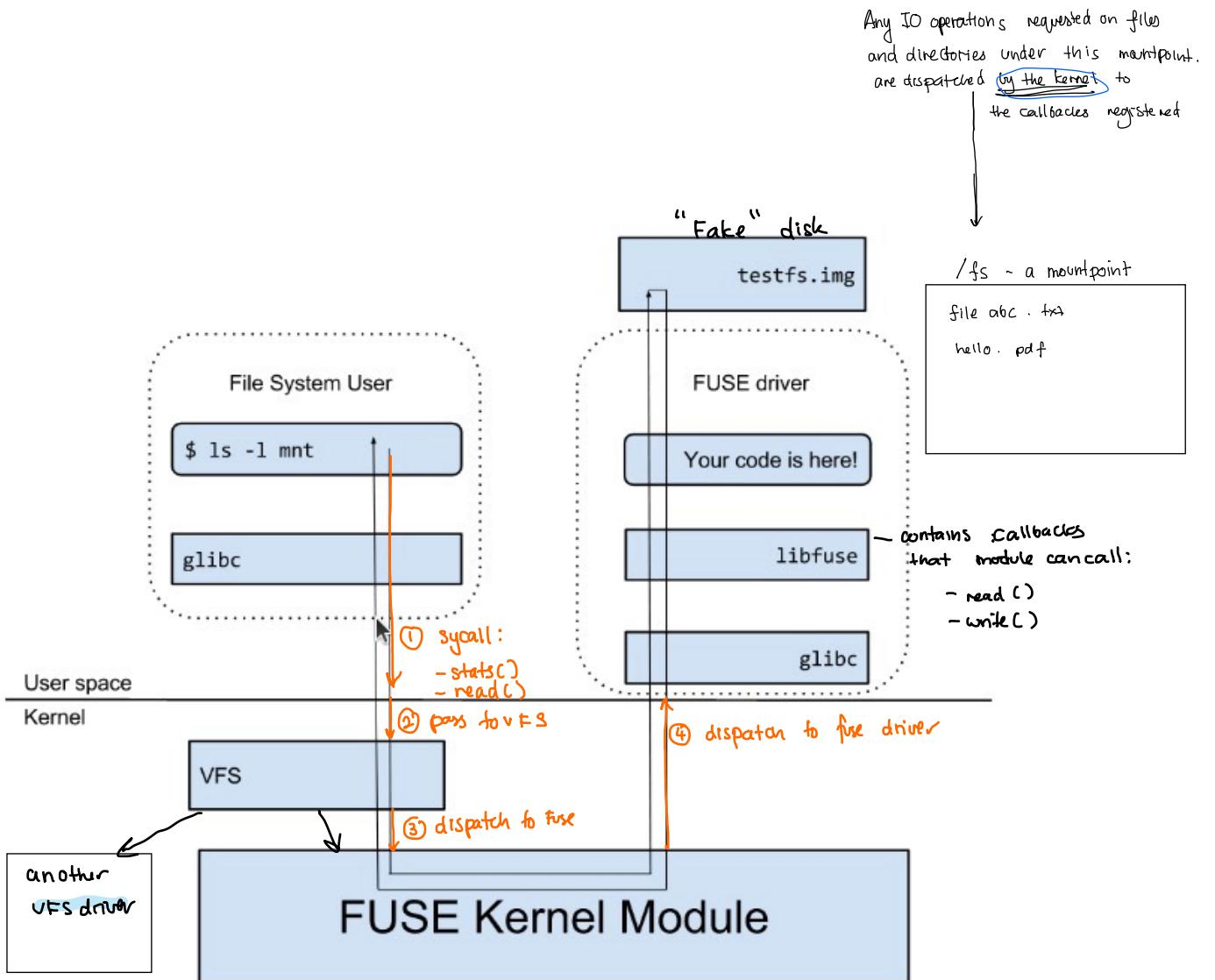
Lab5 Fuse

<https://naizhengtan.github.io/23spring/labtutorials/lab5/>



FUSE allows developers to create and implement file systems without having to write kernel-level code. This approach can simplify development and testing while improving system stability, as errors in the user-space FUSE file system driver are less likely to cause kernel panics or crashes compared to a kernel-level file system implementation.

1. User applications perform file system operations.
2. The VFS layer in the kernel handles the operations and communicates with the appropriate VFS driver. In this case, it's the FUSE kernel module.
3. The FUSE kernel module forwards the requests to the FUSE file system driver in user space.
4. The FUSE file system driver processes the requests and performs the necessary operations.



```

$ mkdir fs      // create a directory to serve as a mount point

$ df fs        // see what file system fs is associated with
Filesystem      1K-blocks    Used Available Use% Mounted on
/dev/sda1       7092728  4536616   2172780  68% /

$ ls fs        // notice, "fs" is empty
① $ ./lab5fuse -image test.img fs    // mount test.img at fs
② $ df fs      // this ends up with an error because you haven't implemented
                  // but notice that fs's file system is different.
df: fs: Operation not supported

$ ls fs        // similarly, "ls" reports an error as well
ls: cannot access 'fs': Operation not supported

$ ls -l        // you will see that fs is in an unknown state
...
d????????? ? ?      ?          ? fs
...

$ fusermount -u fs // now unmount fs

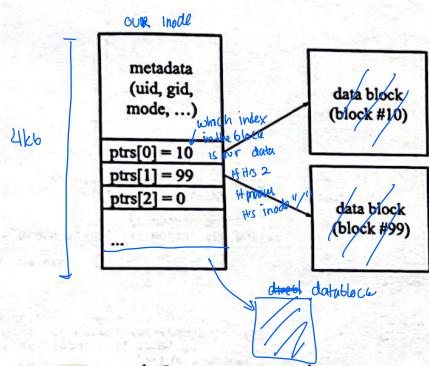
$ ls fs        // and "fs" is an empty dir again

```

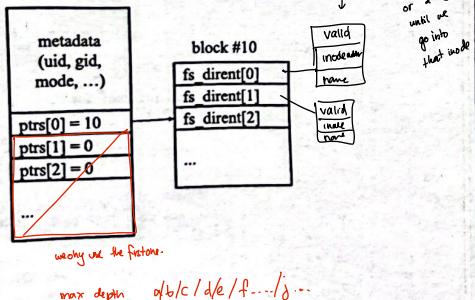
the kernel will dispatch
 - readdir()
 - statfs()
 make to **fuse driver**
 calls that "ls fs"
 - but we haven't implemented

1. fs5600 visualization

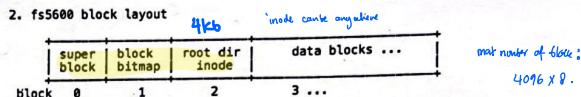
fs5600 file inode:



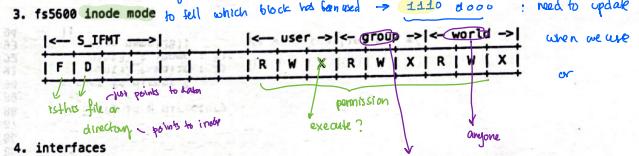
fs5600 directory (also an inode):



2. fs5600 block layout



3. fs5600 inode mode



4. interfaces

```

fs_init - constructor (i.e. put your init code here)
fs_stats - report file system statistics
fs_getattr - get attributes of a file/directory
fs_readdir - enumerate entries in a directory
fs_read - read data from a file
fs_rename - rename a file
fs_chmod - change file permissions

fs_create - Create a new (empty) file
fs_mkdir - create new (empty) directory
fs_unlink - remove a file
fs_rmdir - remove a directory
fs_write - write to a file
fs_truncate - delete the contents of a file
fs_utime - change access and modification times
  
```



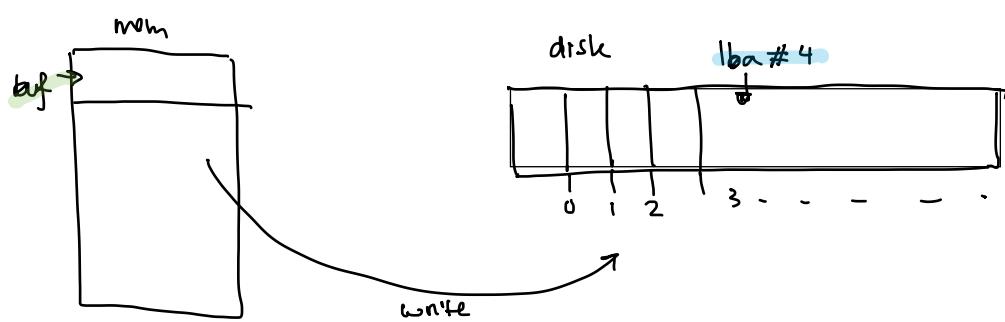
Content Overview

fs5600 disk interface

The fs5600 disk is abstracted as an array of blocks. The size of each block is 4KB (4096 bytes or FS_BLOCK_SIZE). You can read/write disk by two simple interfaces:

```
int block_read(void *buf, int lba, int nblk);
int block_write(void *buf, int lba, int nblk);
```

The `buf` is a pointer to memory where you want to store the data from disk (for reads) or contains the data to disk (for writes). The `lba` is the block id (starting from 0) you want to read/write. The `nblk` is the number of blocks you will read/write.



File system format

The format of fs5600 is as follows:

- the first block (block 0) is the superblock;
- the second block (block 1) is the bitmap block;
- the third block (block 2) is the root directory;
- other blocks are blocks that contain either inodes or data.

| block | 0 | 1 | 2 | 3 | ... |
|-------|-------|-------|----------|-----------------|-----|
| | super | block | root dir | data blocks ... | |

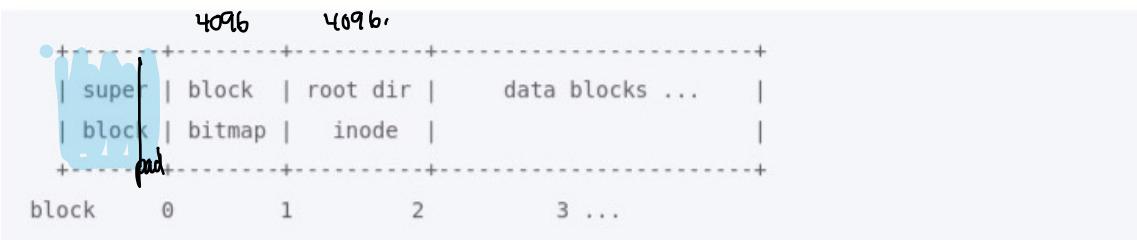
Superblock and block bitmap

Superblock

The superblock is the first block in the file system, and contains the information needed to find the rest of the file system structures. The following C structure (see `fs5600.h`) defines the superblock:

```
struct fsx_superblock {  
    uint32_t magic;           /* 0x30303635 - some magic number */  
    uint32_t disk_size;       /* in 4096-byte blocks */  
    char pad[4088];          /* to make size = 4096 */  
};
```

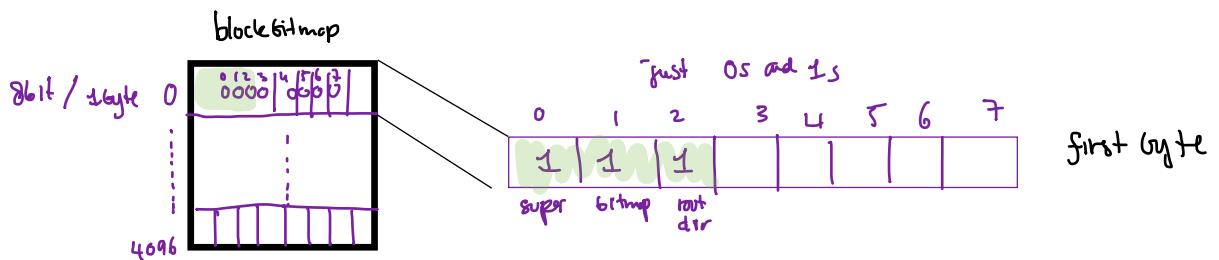
Note that `uint32_t` is a standard C type found in the `<stdint.h>` header file, and refers to an unsigned 32-bit integer. (similarly, `uint16_t`, `int16_t` and `int32_t` are unsigned/signed 16-bit ints and signed 32-bit ints)



Block bitmap

This block bitmap indicates which blocks have been used and which are free. One bit represents one block: the first bit represents the first block (block 0), the second bit represents the second block (block 1), and so forth. If a bit is `0`, the corresponding block is free; otherwise, the block is used.

In `fs5600`, the first three bits are always `1` because they are (1) the super block, (2) the block bitmap itself, and (3) the root dir inode. (the bits for blocks 0, 1 and 2 will be set to 1 when the file system is created, so you don't have to worry about excluding them when you search for a free block.)



3.2 block allocation

Block allocation (with block bitmap)

To allocate blocks from disk, you need to manipulate block bitmap.

The bitmap is cached in a 4KB memory buffer (a global variable `block_bitmap`). You will need keep the cache in sync by writing it back to disk after each block allocation and free.

You're given the following functions to handle the bitmap in memory:

```
int bit_test((void*)map, int i);
void bit_set(unsigned char *map, int i);
void bit_clear(unsigned char *map, int i);
```

A side note: using a single block for the bitmap means that the maximum number of disk blocks fs5600 can track is $32768 (= 4096 * 8)$. Each block is 4KB. Hence the maximum size of an fs5600 is 128MB ($= 32768 * 4KB$).



3.3 Inode

In particular, inode is defined as follows (see also `fs5600.h`):

```

struct fs_inode {
    uint16_t uid;          /* file owner */
    uint16_t gid;          /* group */
    uint32_t mode;         /* type + permissions (see below) */
    uint32_t ctime;        /* last status change time */
    uint32_t mtime;        /* modification time */
    int32_t size;          /* size in bytes */

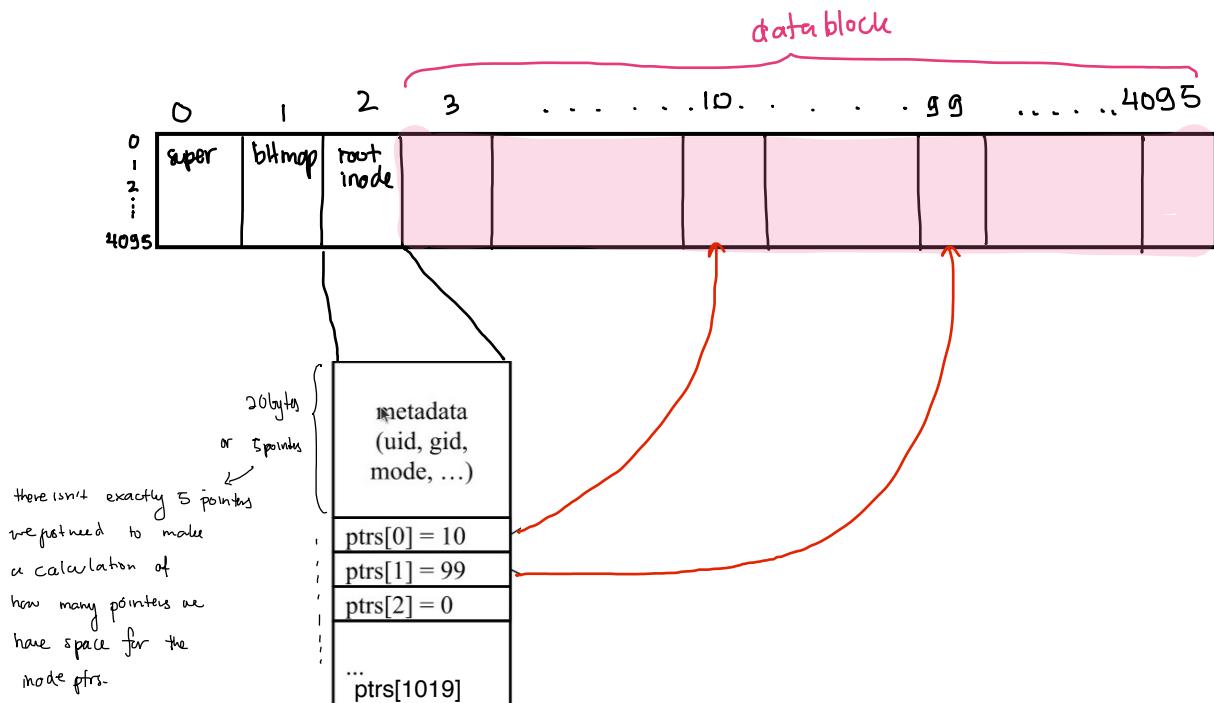
    uint32_t ptrs[FS_BLOCK_SIZE/4 - 5]; /* inode = 4096 bytes */
};

4096 ↓  
4 bytes / pointer  
5 pointers for the other data
total number of pointers

```

Each inode corresponds to a file or directory; in a sense the inode is that file or directory. In fs5600, inode can be uniquely identified by its inode number, which is its block id. Again, **fs5600 uses block ids as inode numbers**. The root directory is always found in inode 2 ; inode 0 (the super block) is invalid and can be used as a `NULL` value.

Hint: a path lookup always begins with inode 2 , the root directory.



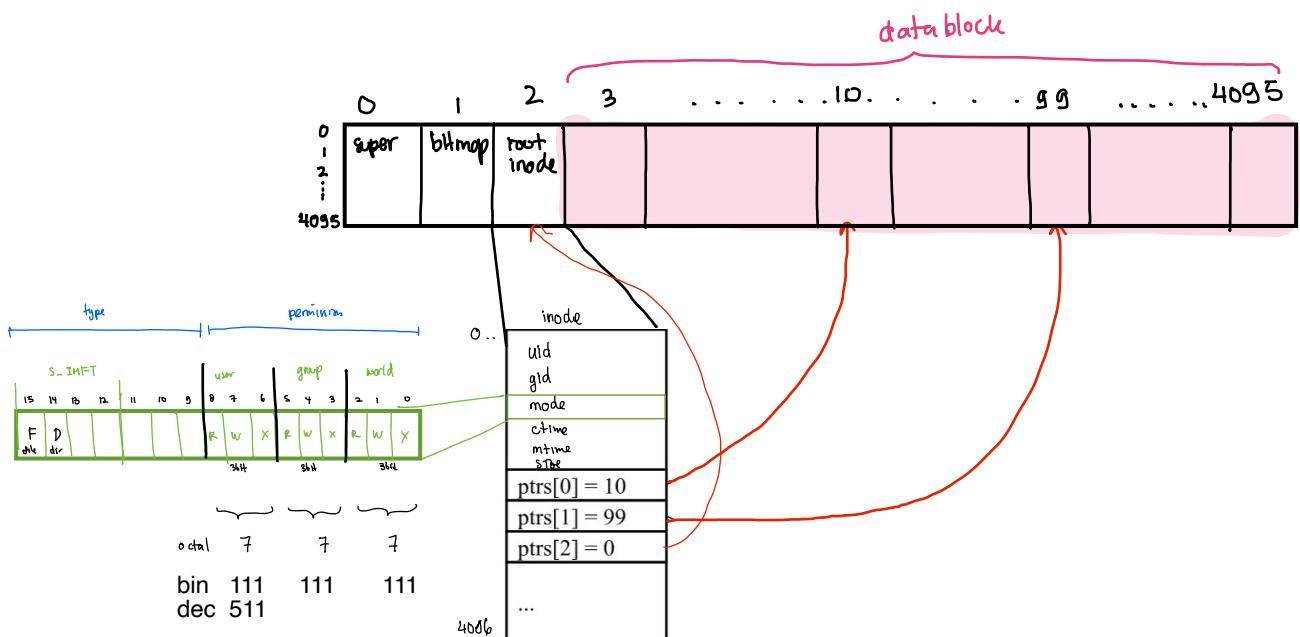
inode's mode

File mode

The FUSE APIs (and Linux internals in general) mash together the concept of object type (file/directory/device/symlink...) and permissions. The result is called the `file mode` (the attribute `mode` in the `fs_inode` above), and looks like this:

```
|<- S_IFMT -->|      |<- user ->|<- group ->|<- world ->|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| F | D |   |   |   | R | W | X | R | W | X | R | W | X |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

00000000 00000000 10000011 10000000



Binary: 10000000 0000 0000
 Grouped: 001 000 000 000 000 000
 Now convert each group of three binary digits to its octal equivalent:
 Octal: 1 0 0 0 0 0

Since it has multiple 3-bit fields, it is commonly displayed in base 8 (octal). For example, permissions allowing `RWX` for everyone (`rwxrwxrwx`) are encoded as `777` (an octal number). Note that, in C, octal numbers are indicated by a leading `0`, e.g., `0777`, so the expression `0777 == 511` is true.

The `F` and `D` bits correspond to `0100000` and `040000`, indicating if the inode is a regular file (`F` bit is set) or a directory (`D` bit is set).

A side note: since the demise of 36-bit computers in the early 70s, Unix permissions are about the only thing in the world that octal is used for.

There are a bunch of macros (in `<sys/stat.h>`) you can use for looking at the mode; you can see the official documentation with the command `man inode`, but the important ones are:

- `S_ISREG(m)` - is it (i.e. the inode with mode `m`) a regular file?
- `S_ISDIR(m)` - is it a directory?
- `S_IFMT` - bitmap for inode type bits.
You can get just the `inode type` with the expression `m & S_IFMT`, and just the `permission bits` with the expression `m & ~S_IFMT`. (note that `~` is bitwise NOT - i.e. `0`s become `1`s and `1`s become `0`s)

①

what is the
type of this
mode



`mode_type = m & S_IFMT`

② set the permission bits

`mode - we want the permission of = 1000 0000 0000 0000`
`S_IFMT = 1110 0000 0000 0000`
`~S_IFMT= 0001 1111 1111 1111`

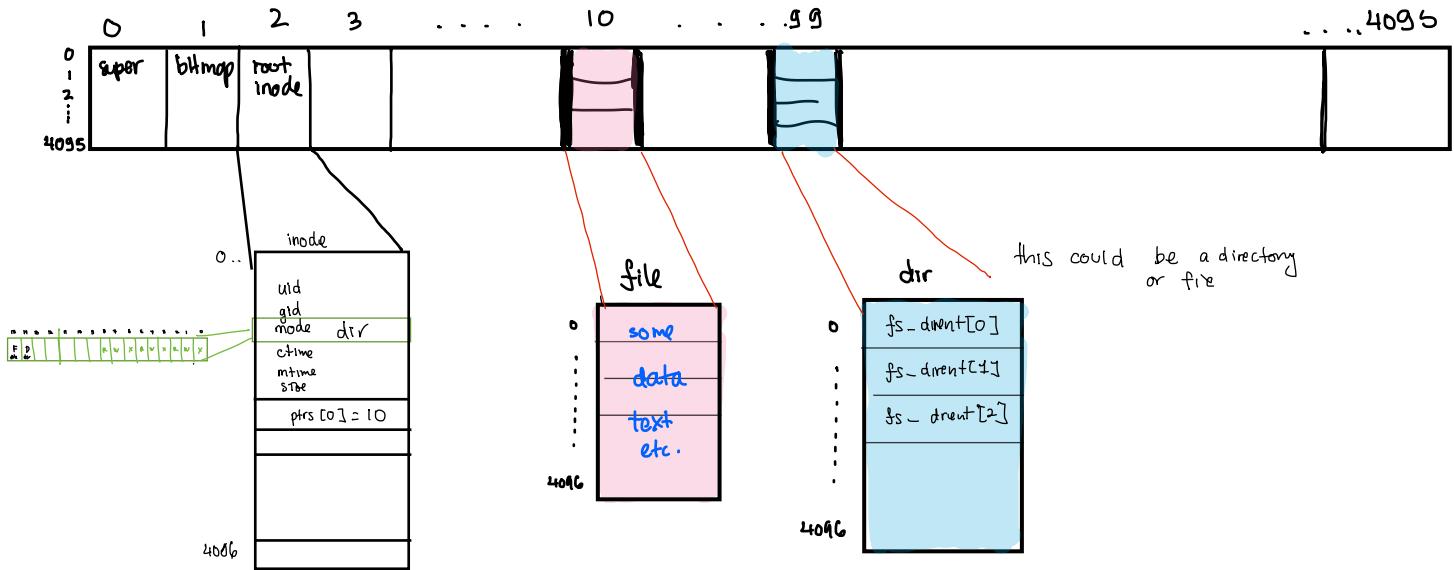
| | | |
|-------|-------|-------------|
| user | group | global |
| 8 | 7 6 | 5 4 3 2 1 0 |
| 0 0 0 | 1 1 1 | 0 0 0 |

`mode - we want the permission of = 1000 0000 0000 0000`
`~S_IFMT= 0001 1111 1111 1111` ✗

`new mode = 0000 000 000 111000`

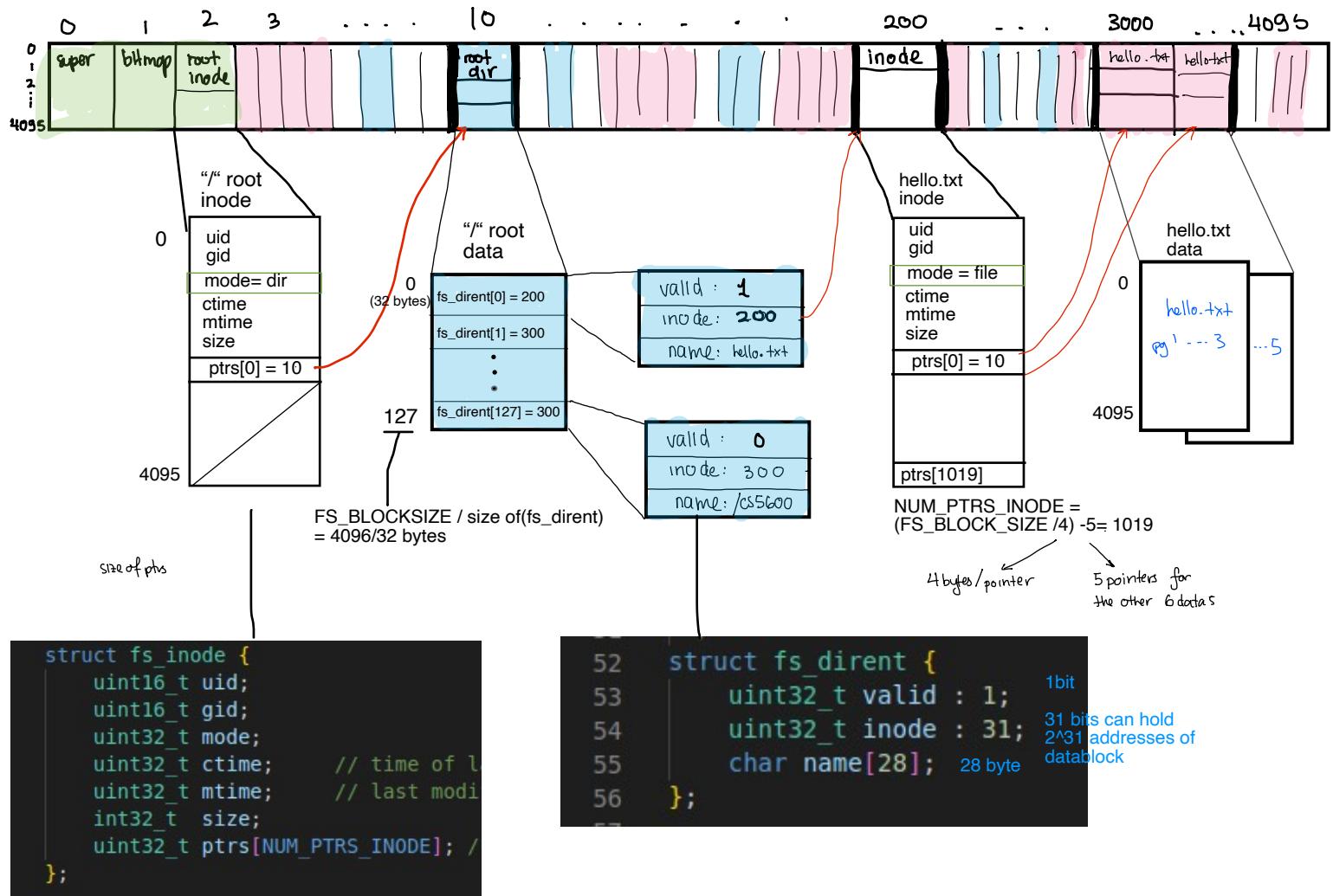
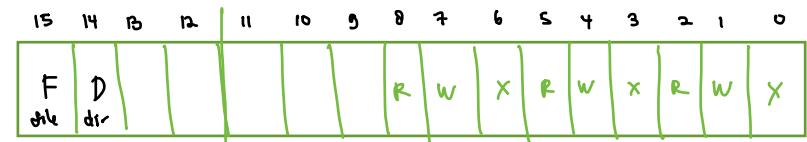
↓
get permission bit only

1. inode pts. can point to a file or a directory



Each directory entry maps a human-readable name (a string) to an inode number (again, an inode number is a block id). It is defined as follows (see also `fs5600.h`):

```
struct fs_dirent {
    uint32_t valid : 1; /* the first bit of uint32_t */
    uint32_t inode : 31; /* the next 31 bits of uint32_t */
    char name[28]; /* with trailing NULL */
};
```



Each `fs_dirent` is 32 bytes, giving 128 (=4096/32) directory entries in one dir. Why? This is a design choice of fs5600: **fs5600 directories always have only one data block**, meaning only the first pointer (`ptrs[0]`) in a dir's inode contains a valid block id; others are invalid (`ptrs[i] == 0` where `i>0`). That said, fs5600's dir has one data block, which is 4KB, and one dirent is 32B, hence a dir only supports a maximum of 128 files.

A directory contains `(name, inode)` pairs, which capture by `fs_dirent`. In particular,

- the `valid` bit indicates if directory entries are being used (1 means yes; 0 means no).
- the `inode` contains a block id on disk that is where the file's inode locates. Note: in fs5600, an inode is a block (4KB in size) and a inode number is the block id on disk.
- `name` is the file name. The maximum name length is 27 bytes, allowing entries to always have a terminating \0 so you can use string functions (e.g., `strcmp`, `strdup`) without any complications.

EX1.1 init the file system

```
310 /* init - this is called once by the FUSE framework at startup.  
311 *  
312 * The function reads the superblock and check if the magic number matches  
313 * FS_MAGIC. It also initializes two global variables:  
314 * "num_blocks" and "block_bitmap".  
315 *  
316 * notes:  
317 * - use "block_read" to read data (if you don't know how it works, read its  
318 * implementation in misc.c)  
319 * - if there is an error, exit(1)  
320 */  
321 void* fs_init(struct fuse_conn_info *conn)  
322 {  
323  
324     struct fs_super sb;  
325     // read super block from disk  
326     if (block_read(&sb, 0, 1) < 0) { exit(1); }  
327  
328     // check if the magic number matches fs5600  
329     if (sb.magic != FS_MAGIC) { exit(1); }  
330  
331     // EXERCISE 1:  
332     // - get number of block and save it in global variable "numb_blocks"  
333     //   (where to know the block number? check fs_super in fs5600.h)  
334     num_blocks = sb.disk_size; // from superbloc  
335  
336     // - read block bitmap to global variable "block_bitmap"  
337     //   (this is a cache in memory; in later exercises, you will need to  
338     //   write it back to disk. When? whenever bitmap gets updated.)  
339     if (block_read(block_bitmap, 1, 1) < 0) {exit(1);}  
340  
341     return NULL;  
342 }
```

EX1.2 Fill in the file system statistics

```
347  /* EXERCISE 1:
348   * statfs - get file system statistics
349   * see 'man 2 statfs' for description of 'struct statvfs'.
350   * Errors - none. Needs to work.
351   */
352  int fs_statfs(const char *path, struct statvfs *st)
353  {
354      /* needs to return the following fields (ignore others):
355       * [DONE] f_bsize = FS_BLOCK_SIZE
356       * [DONE] f_namemax = <whatever your max namelength is>
357       * [TODO] f_blocks = total image - (superblock + block map)
358       * # total data blocks in fs
359       * [TODO] f_bfree = f_blocks - blocks used
360       * [TODO] f_bavail = f_bfree
361       *
362       * it's okay to calculate this dynamically on the rare occasions
363       * when this function is called.
364     */
365
366     st->f_bsize = FS_BLOCK_SIZE;
367     st->f_namemax = 27; // why? see fs5600.h
368     int used_blocks = 0;
369     // for each block
370     for (int i = 0; i < num_blocks; i++) {
371         // iterate through each bit
372         for (int j = 0; j < 8; j++) {
373             if (block_bitmap[i] & (1 << j)) {
374                 used_blocks++;
375             }
376         }
377     }
378     st->f_blocks = num_blocks;
379     st->f_bfree = st->f_blocks - used_blocks;
380     st->f_bavail = st->f_bfree;
381     return 0;
382 }
```

```
    struct statfs {
    __fsword_t f_type;      /* Type of filesystem (see below) */
    __fsword_t f_bsize;     /* Optimal transfer block size */
    fsblkcnt_t f_blocks;   /* Total data blocks in filesystem */
    fsblkcnt_t f_bfree;    /* Free blocks in filesystem */
    fsblkcnt_t f_bavail;   /* Free blocks available to
                                unprivileged user */
    fsfilcnt_t f_files;    /* Total inodes in filesystem */
    fsfilcnt_t f_ffree;    /* Free inodes in filesystem */
    fsid_t     f_fsid;     /* Filesystem ID */
    __fsword_t f_namelen;  /* Maximum length of filenames */
    __fsword_t f_frsize;   /* Fragment size (since Linux 2.6) */
    __fsword_t f_flags;    /* Mount flags of filesystem
                                (since Linux 2.6.36) */
    __fsword_t f_spare[xxx];
                           /* Padding bytes reserved for future use */
};
```

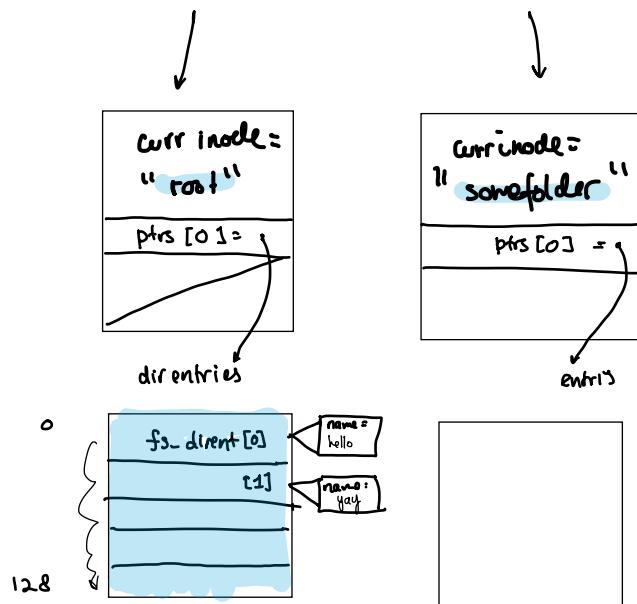
Ex 2:1

Part 1 helper: path2inum(path):

at every token iteration, the curr node has to be a dir type.
otherwise it cannot find our token

depth = 2

tokens = ["file.8kt", "file.0"]
find in find in



test1.c:223:F:read_mostly:path_errs:0: Assertion 'rv == -bad1[i].err' failed: rv == -2, -bad1[i].err == -20

/file.8k+/file.0 should give -20 ENOTDIR
but mins shows ENOENT

Test cases :

```
# path uid gid mode ctime mtime
"/", 0, 0, 040777, 4096, 1565283152, 1565283167
"/file.1k", 500, 500, 0100666, 1000, 1565283152, 1565283152
"/file.10", 500, 500, 0100666, 10, 1565283152, 1565283167
"/dir-with-long-name", 0, 0, 040777, 4096, 1565283152, 1565283167
"/dir-with-long-name/file.12k+", 0, 500, 0100666, 12289, 1565283152, 1565283167
"/dir2", 500, 500, 040777, 8192, 1565283152, 1565283167
"/dir2/twenty-seven-byte-file-name", 500, 500, 0100666, 1000, 1565283152, 1565283167
"/dir2/file.4k+", 500, 500, 0100777, 4098, 1565283152, 1565283167
"/dir3", 0, 500, 040777, 4096, 1565283152, 1565283167
"/dir3/subdir", 0, 500, 040777, 4096, 1565283152, 1565283167
"/dir3/subdir/file.4k-", 500, 500, 0100666, 4095, 1565283152, 1565283167
"/dir3/subdir/file.8k-", 500, 500, 0100666, 8190, 1565283152, 1565283167
"/dir3/subdir/file.12k", 500, 500, 0100666, 12288, 1565283152, 1565283167
"/dir3/file.12k-", 0, 500, 0100777, 12287, 1565283152, 1565283167
"/file.8k+", 500, 500, 0100666, 8195, 1565283152, 1565283167
```

helper we can use :

```
Because tests of the above form are common, additional macros
are defined by POSIX to allow the test of the file type in
st_mode to be written more concisely:
I
S_ISREG(m)  is it a regular file?
S_ISDIR(m)  directory?
S_ISCHR(m)  character device?
S_ISBLK(m)  block device?
S_ISFIFO(m) FIFO (named pipe)?
S_ISLNK(m)  symbolic link? (Not in POSIX.1-1996.)
S_ISSOCK(m) socket? (Not in POSIX.1-1996.)
```

- ENOENT Ifile/dir not found
- EIO -if block_read or block_write fail
- ENOMEM -out of memory
- ENOTDIR -not a directory
- EISDIR -is a directory
- EINVAL -invalid parameter (see comments in fs5600.c)
- EOPNOTSUPP -not implemented yet

Ex 2:2

helper: inode2stat(stat sb, fs_inode in, inode_num):

copy the info in an **inode** to **struct stat**

```
struct stat {  
    ino_t      st_ino;          // Inode number  
    mode_t     st_mode;         // File type and mode  
    nlink_t    st_nlink;        // Number of hard links (set to 1)  
    uid_t      st_uid;          // User ID of owner  
    gid_t      st_gid;          // Group ID of owner  
    off_t      st_size;         // Total size, in bytes  
    blkcnt_t   st_blocks;       // Number of blocks allocated  
                                // (note: block size is FS_BLOCK_SIZE;  
                                // and this number is an int which should be round up)  
  
    struct timespec st_atim;   // Time of last access (same as st_mtime)  
    struct timespec st_mtim;   // Time of last modification  
    struct timespec st_ctim;   // Time of last status change  
};
```



```
struct fs_inode {  
    uint16_t uid;  
    uint16_t gid;  
    uint32_t mode;  
    uint32_t ctime;    // time of l  
    uint32_t mtime;    // last modi  
    int32_t  size;  
    uint32_t ptrs[NUM_PTRS_INODE]; /  
};
```

Ex 2.3: fs_getattr()

```
int fs_getattr(const char *path, struct stat *sb)
{
    printf("\n>>>getattr");

    // 1. get the inode num and inode
    int inodenum = path2inum(path);

    // check if inodenum returns error
    if (inodenum < 0) {
        return inodenum; // return the error code
    }

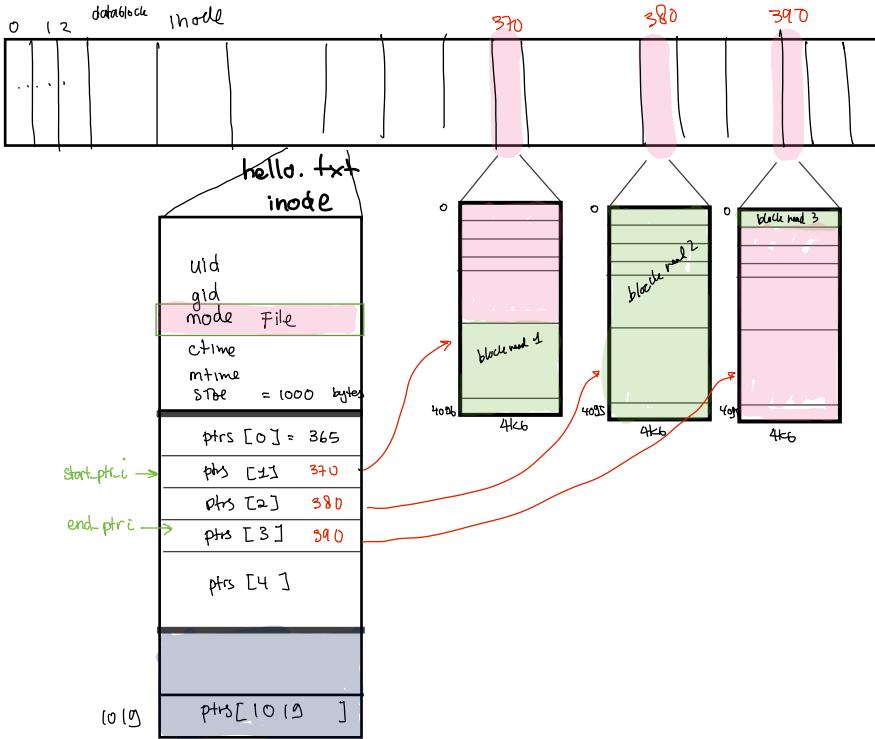
    // get the inode
    struct fs_inode curr_inode;
    if (block_read(&curr_inode, inodenum, 1) < 0) {
        free(path);
        return -EIO;
    }
    // 2. copy inodes info to struct stat
    inode2stat(sb, &curr_inode, inodenum);
    print_node_info(inodenum, curr_inode);
    print_stat(sb);

    // 3. return 0
    return 0;
}
```

Ex 2.4: fs_readdir() get all the entries in directories. we will use ptr (the second argument) to fill this all the info

```
464 int fs_readdir(const char *path, void *ptr, fuse_fill_dir_t filler,
465 |   off_t offset, struct fuse_file_info *fi)
466 {
467 |   // printf("\n>>>readdir\n");
468 |   // 1. get the inode num and inode
469 |   // int dir_inodenum = path2inum(path);
470 |   char *_path = strdup(path);
471 |   int dir_inodenum = path2inum(_path);
472 |
473 |   // 2. check if inodenum returns error
474 |   if (dir_inodenum < 0) {
475 |       return dir_inodenum; // return the error code
476 |   }
477 |
478 |   // 3. get the inode
479 |   struct fs_inode dir_inode;
480 |   if (block_read(&dir_inode, dir_inodenum, 1) < 0) {
481 |       free(_path);
482 |       return -EIO;
483 |   }
484 |   // 4. make sure this is a dir's inode
485 |   if (!S_ISDIR(dir_inode.mode)) {
486 |       free(_path);
487 |       return -ENOTDIR;
488 |   }
489 |
490 |   // 5. copy the dir_entries to mem
491 |   // get all its entries from disk (4096/32B = 128 entries) to memory
492 |   int DIR_ENTRY_NUM = FS_BLOCK_SIZE / sizeof(struct fs_dirent);
493 |   struct fs_dirent dir_entries[DIR_ENTRY_NUM];
494 |
495 |   if (block_read(dir_entries, dir_inode.ptrs[0], 1) < 0) {
496 |       free(_path);
497 |       return -EIO;
498 |   }
499 |
500 |   // 6. iterate through each entry
501 |   for (int dir_entry_i = 0; dir_entry_i < DIR_ENTRY_NUM; dir_entry_i++) {
502 |       if (dir_entries[dir_entry_i].valid == 1) {
503 |           // 1. get the name of this entry
504 |           char* entry_name = dir_entries[dir_entry_i].name;
505 |           uint32_t entry_inodenum = dir_entries[dir_entry_i].inode;
506 |
507 |           // 2. get the statbuf of this entry
508 |           // - get inode
509 |           struct fs_inode entry_inode;
510 |           if (block_read(&entry_inode, entry_inodenum, 1) < 0) {
511 |               free(_path);
512 |               return -EIO;
513 |           }
514 |
515 |           // - use the inode to get statbuf
516 |           struct stat entry_statbuf;
517 |           inode2stat(&entry_statbuf, &entry_inode, entry_inodenum);
518 |
519 |           // 3. fill
520 |           filler(ptr, entry_name, &entry_statbuf, 0);
521 |
522 |       }
523 |   }
524 |   // printf("finish iterating entries>>>>>>\n");
525 |   // return success
526 |   return 0;
527 }
528 }
```

EX3.1 : fs_read()



① validate block read file inode

`ptrs[0] =`

② adjust length

$$\text{total \# phs} = \lceil 1000 / 4096 \rceil$$

we want to start from `start_ith_bytes = 986th bytes` of the data block

we want to read `bytes_num_to_read = 17 bytes`

$$\text{so the } \text{end_ith_bytes} = 986$$

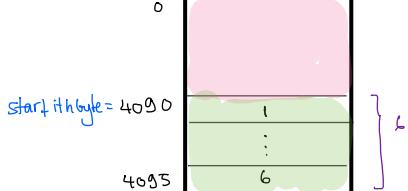
$$= 1003 \text{ too big!}$$

$$\text{adjust length: } \text{bytes_num_to_read} = 1000 - 986 = 14 \text{ bytes}$$

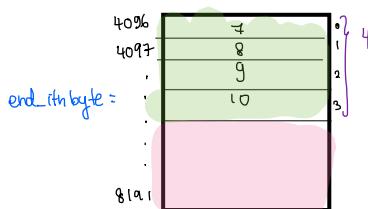
③ get the block number to read

$$\text{start_ptr_i} = \frac{\text{start_ith_bytes}}{4096}$$

$$\text{end_ptr_i} = \frac{(\text{start_ith_bytes} + \text{bytes_num_to_read} - 1)}{4096}$$



`ptrs[1] =`



⑤ read block

block 1

ptrs[2]

0/8192

block_start_i
start_ih byte offset

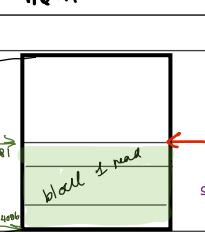
4082/12,274

block_end_i
end_ih byte

4095/12,287

4096

mem



mem-block = x

size = 14

5.4 memory (buf, src, 14)

5.5 increment buffer by size 14

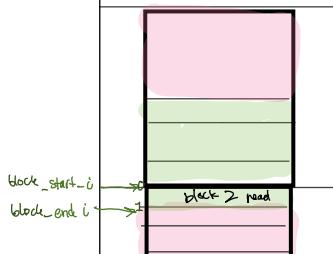
block 2

ptrs[3]

block_start_i = 0 / 12,228
block_end_i = 1 / 12,289
end_ih byte

15

mem



mem-block 2
src1 = x + 0
= x

5.3 len = 1

5.4 memory (buf, src, 1)

5.5 buf += (len)

⑥ read block from

start_ptr_i ~ end_ptr_i :

1 read block i

2 get block_start_i
and
block_end_i

based on nth byte

3 get the length of this block

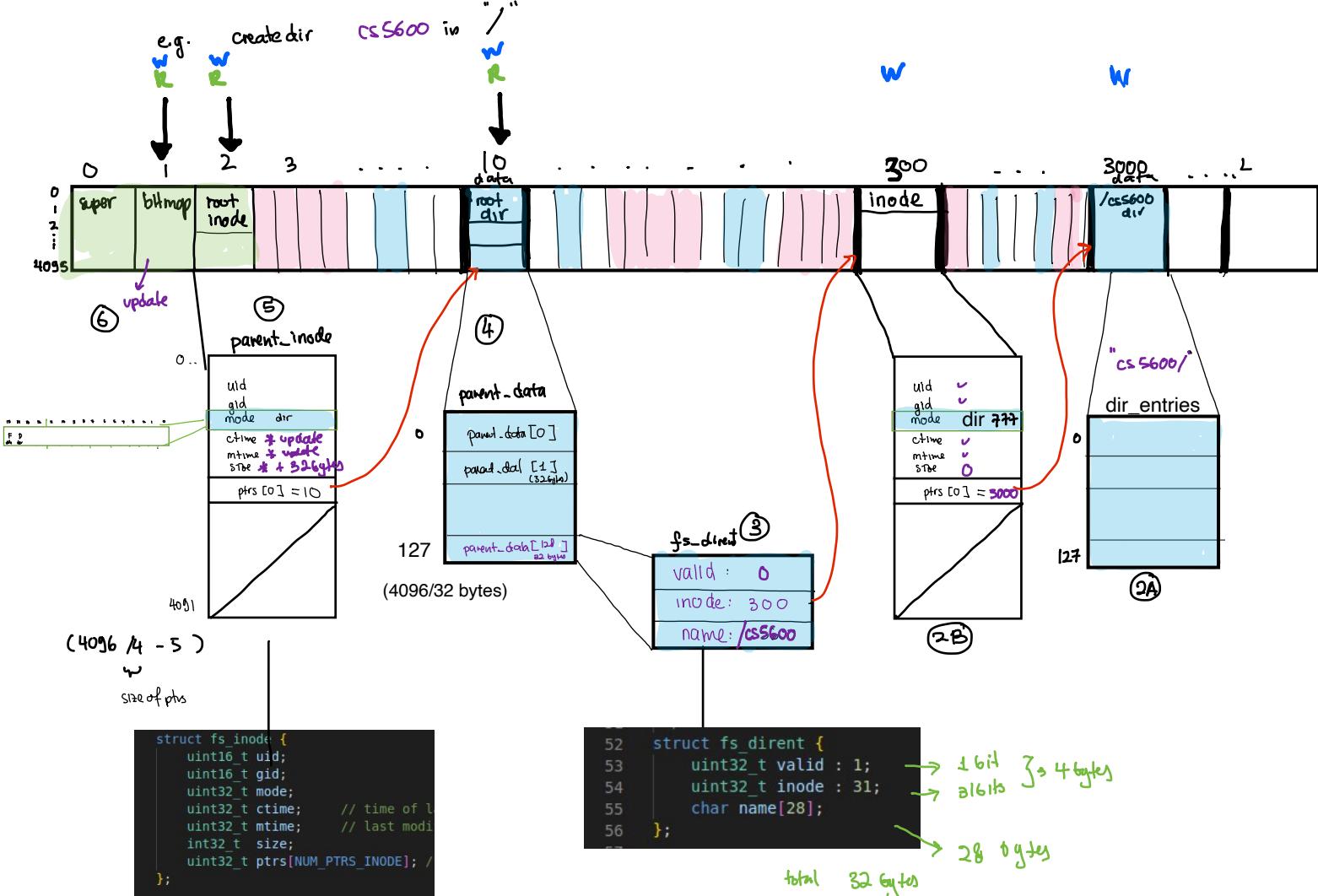
4. copy

5 increment buf ptr with length

. block_write() hints:

1. bitmap
2. inode # 10 ↳ a new inode
3. datablock# 11 ↳ data (empty)
4. write to data block of root "/"
directly
5. modify the inode 2:

Ex 4.1 fs_mkdir



part①: validate

part①: find 2 empty blocks from bitmap

1. read_block (bit_block, 2)

2. get 2 bits that are 0 :

1 for new dir's inode num
1 for new dir's datablock num

Part ② inode & data for cs5600

(A) • init empty data

(B) • inode → uid → ctime
→ gid → mtime
→ mode → size → 0
→ ptrs[0] = new dir's datablock num

Part ③ fill in the fs_dirent of cs5600

| |
|------------------------------------|
| valid = 1 |
| inode = <u>new dir's inode num</u> |
| name = cs5600 |

Part ④ data for parent update

1. read_block the parent node (to mem)
par_dir_entries
2. read_block the parent data (to mem)
3. find "invalid" entry in new
→ assign cs5600 fs_dirent to this index

Part ⑤ node of parent update

Part ⑥ update bitmap set to 1

on newdir_inode_num
newdir_datablock_num

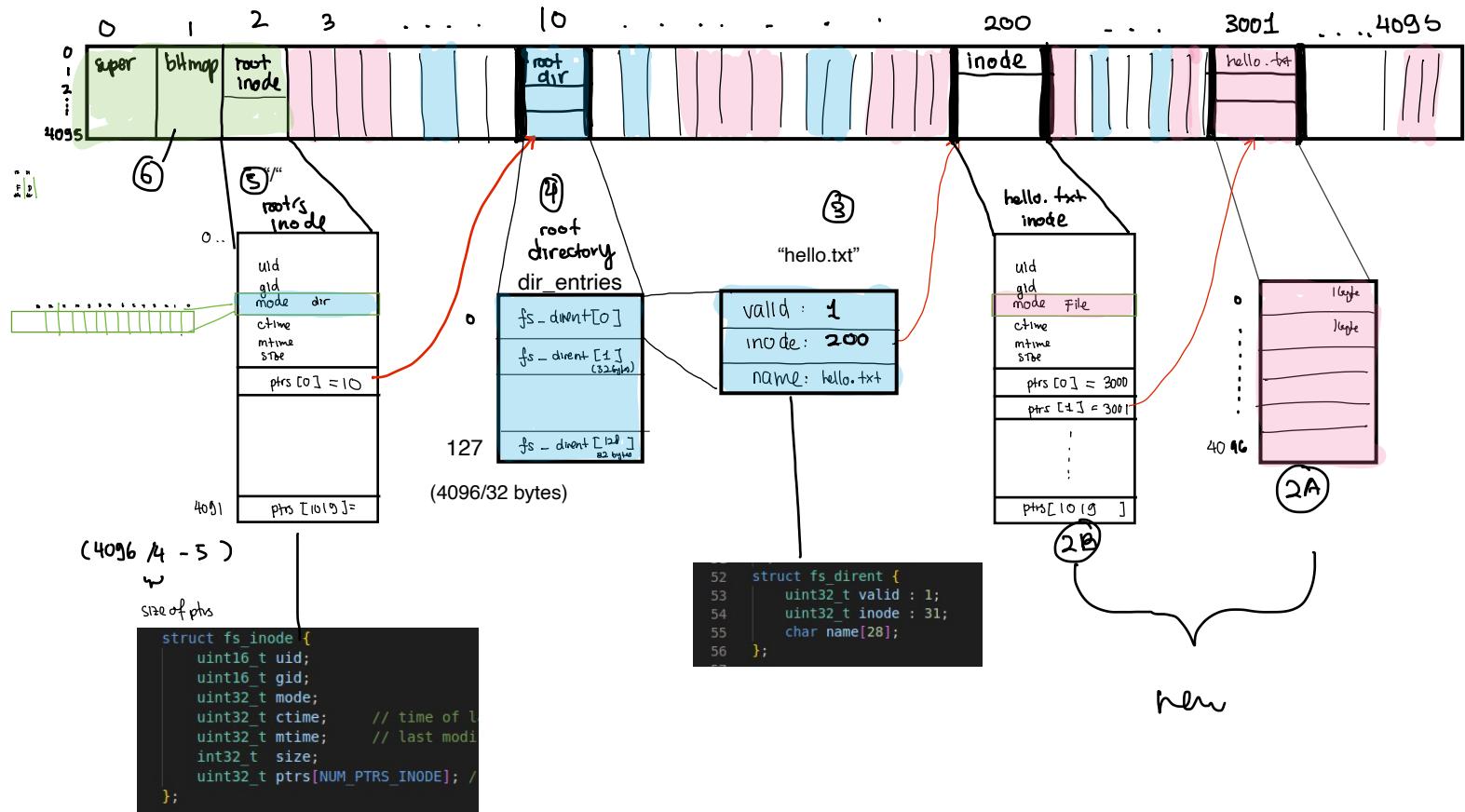
Part ⑦

write block: parent inode

parent data (direnty.)
cs5600 inode
cs5600 data (direnty.)
bitmap

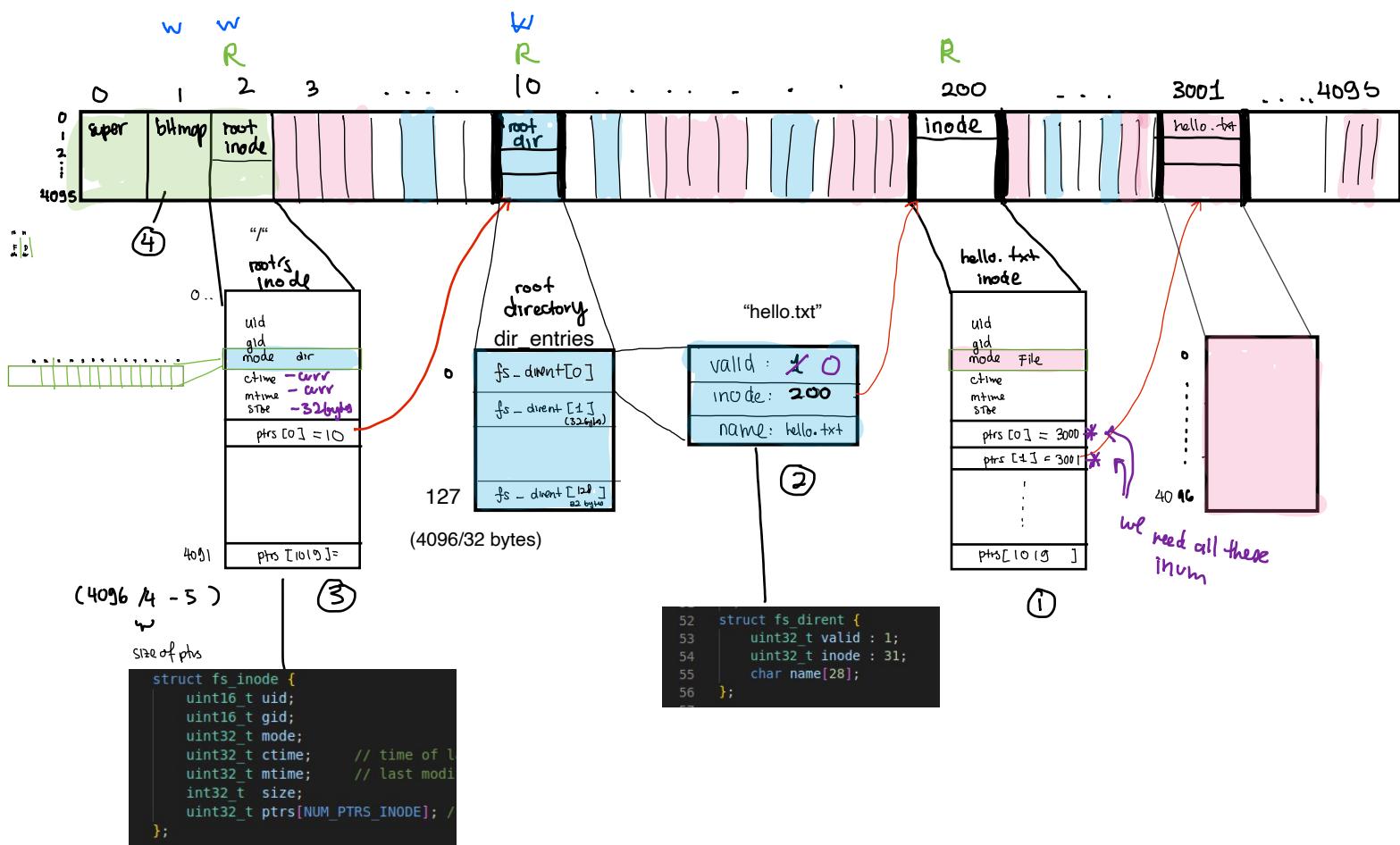
Ex 4.2 fs create

create file e.g. hello.txt



part o validate

Ex 5.1 fs_unlink (delete a file)



part 0 validate and

make sure we have **block_read** these:

- parent_inode
- parent_data
- file_inode

part 1 get info from file's mode

- store all the inum in int[3] inums_delete[i]
- free file

part 2 update parent's datablock

in data block of parent, find the filename,

- set valid to 0
- add the inum to inums-del
- free file

part 3 update parent's inode

update ctime and mtime, clear size by -32bytes

part 4: update block bitmap

part 5: blockwrite

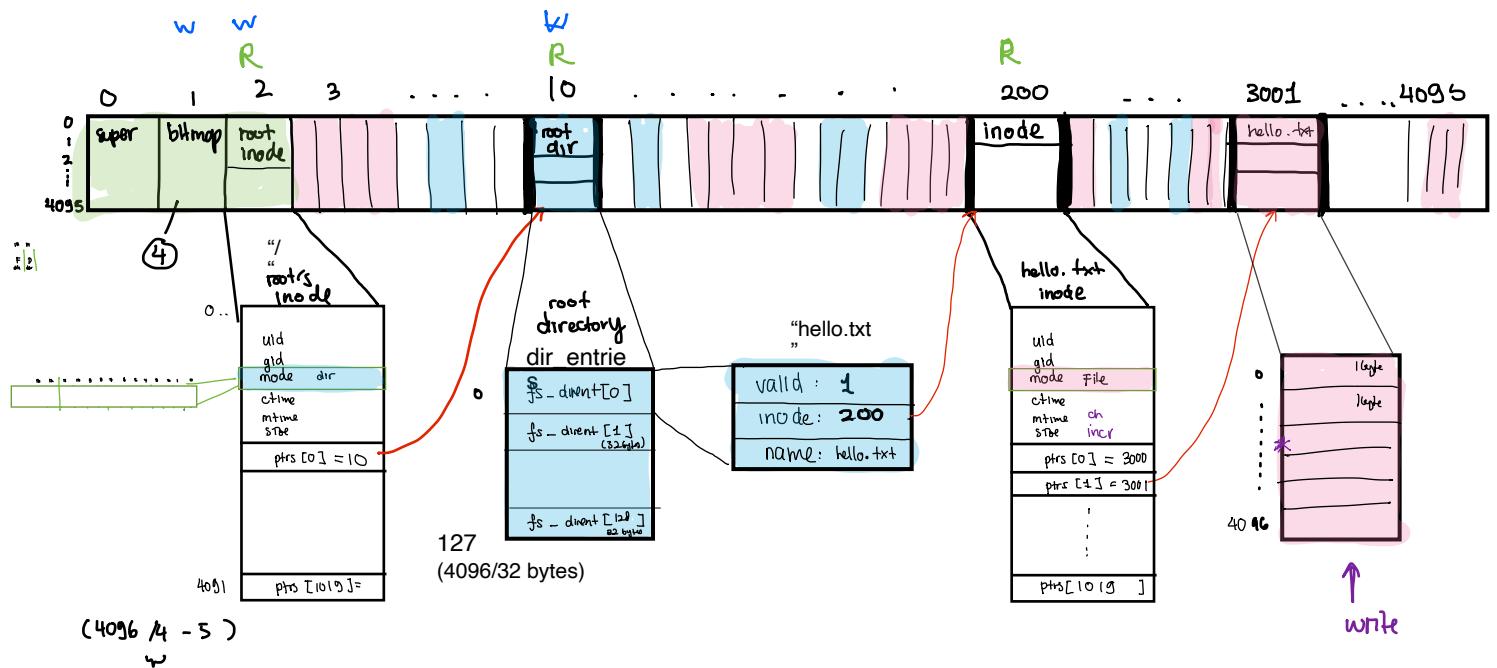
blockwrite : bitmap
parent_inode
parent_data

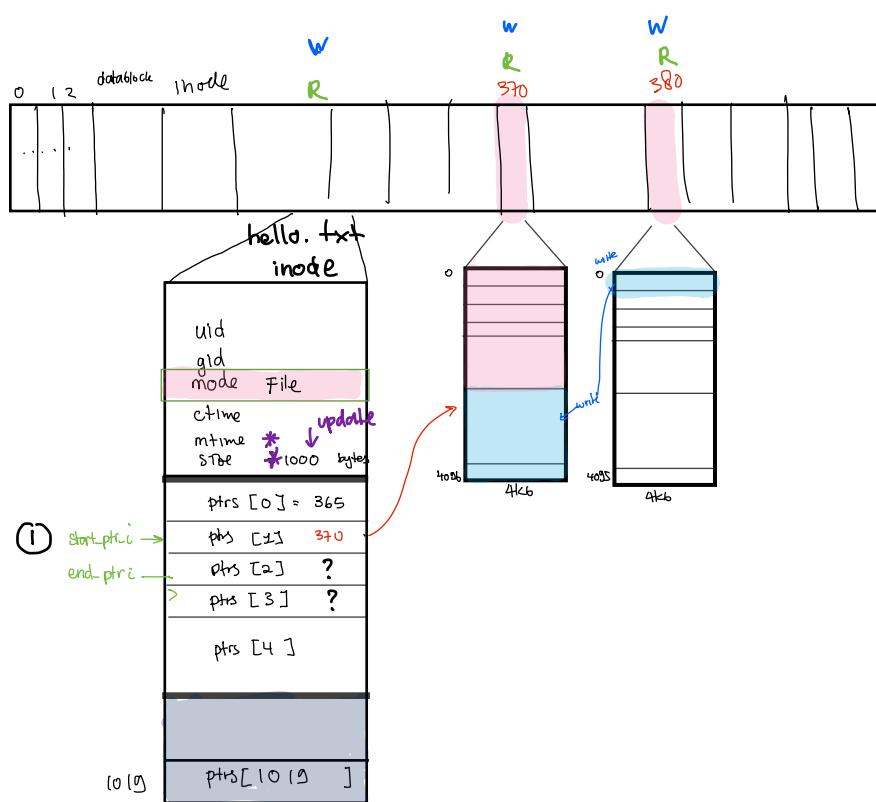
EXERCISE 5 implement `fs_unlink` and `fs_rmdir`.

- implement `fs_unlink`, removing a file
- `fs_unlink` errors:
 - bad path `/a/b/c - b` doesn't exist (`-ENOENT`)
 - bad path `/a/b/c - b` isn't directory (`-ENOTDIR`)
 - bad path `/a/b/c - c` doesn't exist (`-ENOENT`)
 - bad path `/a/b/c - c` is directory (`-EISDIR`)
- implement `fs_rmdir`, removing a directory
- `fs_rmdir` errors:
 - bad path `/a/b/c - b` doesn't exist (`-ENOENT`)
 - bad path `/a/b/c - b` isn't directory (`-ENOTDIR`)
 - bad path `/a/b/c - c` doesn't exist (`-ENOENT`)
 - bad path `/a/b/c - c` is file (`-ENOTDIR`)
 - directory not empty (`-ENOTEMPTY`)
- after implementing, you should pass another 10 cases:

```
$ make testb  
***  
51%: Checks: 27, Failures: 13, Errors: 0
```

Ex 6.1 fs write (to a file)





Part ① validate

block read
file inode
file data

- ① file - num not found (enonent) 1. num
- ② file - inode is a dir (ersdir) 2. inode
- ③ offset = start - 4h bytes > curr file length (EINVAL) 3. off - file bytes
max file size
- ④ bytes num to write + file inode. size > (1019 * 4kb) + len
ENOSPC

But ① got start_ptr_i and end_ptr_i we can read all the required block number

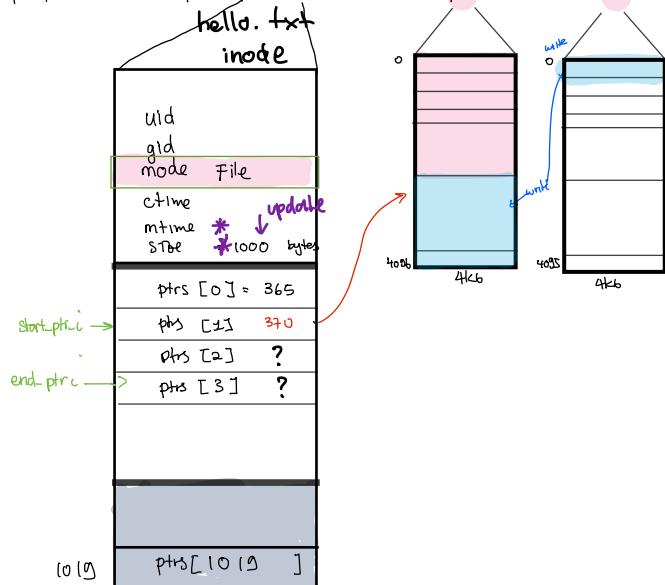
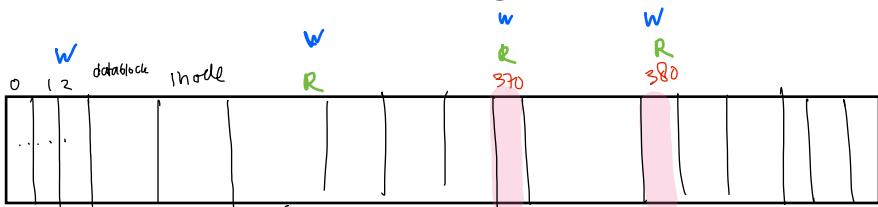
$$\text{end_ith_bytes} = \text{start_ith_bytes} + \text{bytes.num_to_read}$$

(ceil.)

$$\text{start_ptr_i} = \frac{\text{start_ith_bytes}}{4096} = 1$$

$$\text{end_ptr_i} = \left(\frac{\text{end_ith_bytes}}{4096} - 1 \right) / 4096 = 2$$

Part ② check block by block if data block already exist



X

if Data_inum is valid check if it's bitset to 0 (which means we can block-read this)

```
int data_inum = file_inode.ptrs[i]
```

//case A: already exist, just load to mem

```
if (data_inum >= 3 and data_inum < num_blocks)
    and bitset (block_bitmp, data_inum) {
        mem      disk
        dst      src
        block read (mem_full_block, data_inum, 1)
```

3

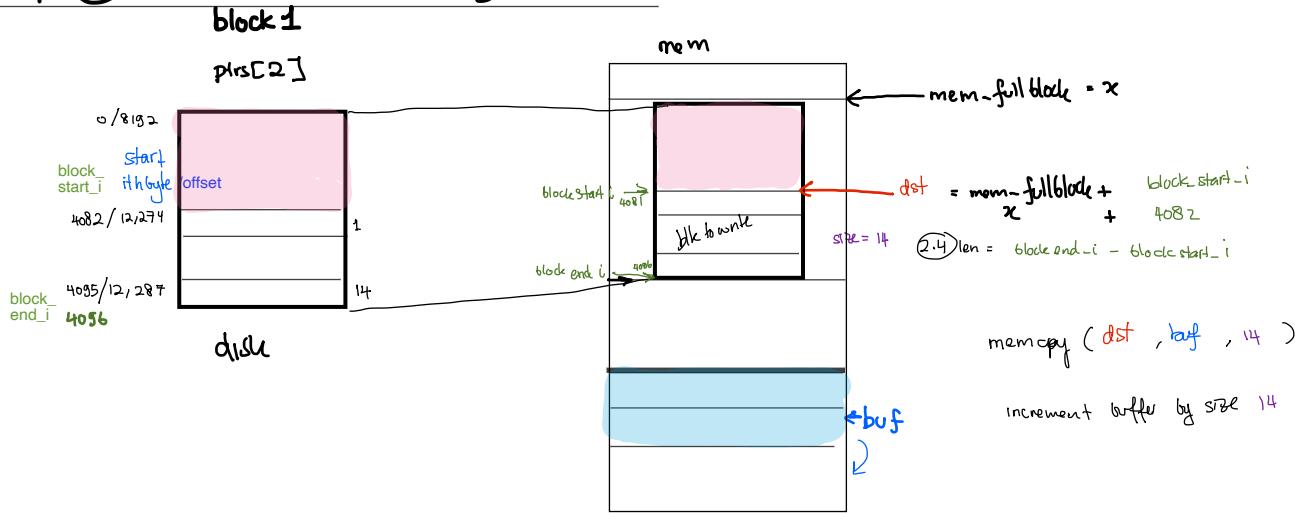
//case B: not yet exist, init new datablock

- int new_datablocknum = alloc_blk()

- insert this to file_inode.ptrs

- init memblock in memory.

Part ② write block by block



2 get the `block_start_i` to know when to start writing.

```
if this is the first block
    ptr_i == start_ptr_i
    block_start_i = start_i-th byte % FS_BLOCK_SIZE
else: 0
```

3. get the `block_end_i` (exclusive) to know when to finish writing:

```
if this is the last block and the last byte is to write is not the
    ptr_i == end_ptr_i     end_i-th byte % FS_BLOCK_SIZE != 0
```

```
block_end_i < end_i-th byte % FS_BLOCK_SIZE
else:
    FS_BLOCK_SIZE
```

4. get the length of data to write

$$\text{len_write_pblock} = \text{block_start_i} - \text{block_end_i}$$

5. get dst address in memblock full where to write

$$\text{dst} = \text{mem_full_block} + \text{block_start_i}$$

6. block write this data from buffer mem_full_block to
 $\text{no_file_inode_ptrs[i]}$

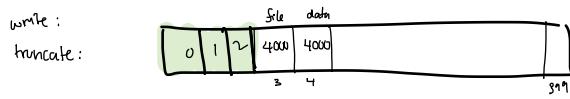
7. incr. buf

Part 3 update file's size and

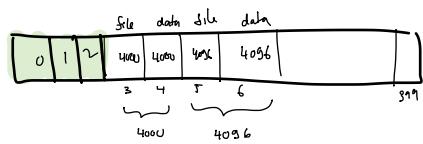
fs_truncate:



file-2000



file-4096



Chall. caesar

1. get_caesar_key() will return the key if file exist.

2. fs-init():

```
// to test:  
// fs-create ("key");  
// fs-write ("2");  
// fs-create ("abc.txt");  
// fs-write ("abc");
```

2. fs-read(path, wif, len, offset, fi):

// 1. init caesar key

```
int caesar-key = 0;
```

if path is not "key":

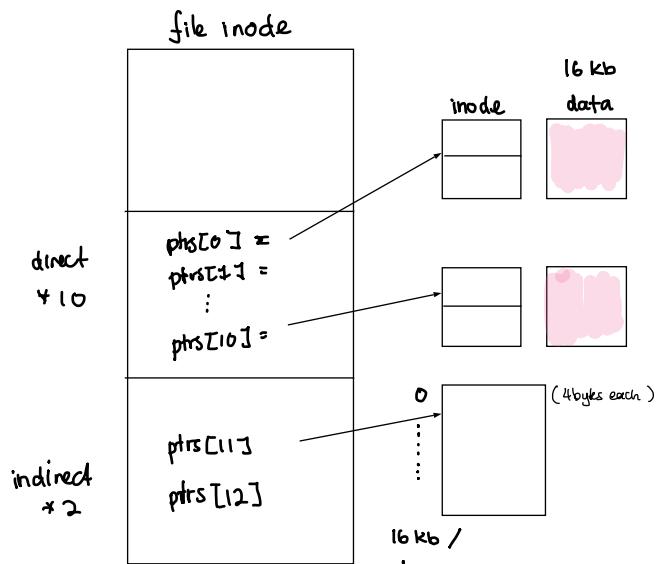
```
caesar-key = get_caesar-key()
```

// 2. when we read

Server 5

- `paxosReplica. putRequest ()`
- `putRequest (key, val)`:
 - 1 → get leader r : `leader Verification ()`
 - if no leader
 - start `election()`.
 - `receive Election message (known Max ID)`:
 - case B → if leader not established &
and I'm not the leader.
ask the next node
 - `receive Election message (known Max ID)`:
 - case A → I am the known MaxID:
 - I'm the leader

2. `this.leaderID` should be 5



Should be able to point to all
 2^{20} of 16 kb
means it should be able
to store 2^{20} addresses.

