

# Exercise 5.1

请根据kern/syscall\_all.c中的说明，完成sys\_write\_dev函数以及sys\_read\_dev函数的实现。

编写这两个系统调用时需要注意物理地址与内核虚拟地址之间的转换。

同时还要检查物理地址的有效性，在实验中允许访问的地址范围为：

console:[0x10000000, 0x10000020), disk:[0x13000000, 0x13004200), rtc:[0x15000000, 0x15000200)，当出现越界时，应返回指定的错误码。

```
1  /* Overview:
2   * This function is used to write data at 'va' with length 'len' to a device
   physical address
3   * 'pa'. Remember to check the validity of 'va' and 'pa' (see Hint below);
4   *
5   * 'va' is the starting address of source data, 'len' is the
6   * length of data (in bytes), 'pa' is the physical address of
7   * the device (maybe with a offset).
8   *
9   * Post-Condition:
10  * Data within [va, va+len) is copied to the physical address 'pa'.
11  * Return 0 on success.
12  * Return -E_INVAL on bad address.
13  *
14  * Hint: Use the unmapped and uncached segment in kernel address space (KSEG1)
   to perform MMIO.
15  * Hint: You can use 'is_illegal_va_range' to validate 'va'.
16  * Hint: You MUST use 'memcpy' to copy data after checking the validity.
17  *
18  * All valid device and their physical address ranges:
19  * * -----*
20  * | device | start addr | length |
21  * * -----+-----+-----*
22  * | console | 0x10000000 | 0x20 | (dev_cons.h)
23  * | IDE disk | 0x13000000 | 0x4200 | (dev_disk.h)
24  * | rtc      | 0x15000000 | 0x200 | (dev_rtc.h)
25  * * -----*
26  */
27 int sys_write_dev(u_int va, u_int pa, u_int len) {
28     /* Exercise 5.1: Your code here. (1/2) */
29     if(is_illegal_va_range(va, len)) {
30         return -E_INVAL;
31     }
32
33
34     if ((0x10000000 <= pa && pa + len <= 0x10000020) ||
35         (0x13000000 <= pa && pa + len <= 0x13004200) ||
36         (0x15000000 <= pa && pa + len <= 0x15000200)) {
37         memcpy((void *) (pa + KSEG1), (void *) va, len); //va的数据复制到pa; 注意
   pa + KSEG1转内核地址
38         return 0;
```

```

39     }
40
41     return -E_INVALID;
42 }

```

```

1  /* Overview:
2   * This function is used to read data from a device physical address.
3   * Remember to check the validity of addresses (same as in 'sys_write_dev').
4   *
5   * Post-Condition:
6   * Data at 'pa' is copied from device to [va, va+len).
7   * Return 0 on success.
8   * Return -E_INVALID on bad address.
9   *
10  * Hint: You MUST use 'memcpy' to copy data after checking the validity.
11  */
12  int sys_read_dev(u_int va, u_int pa, u_int len) {
13      /* Exercise 5.1: Your code here. (2/2) */
14      if(is_illegal_va_range(va, len)) {
15          return -E_INVALID;
16      }
17
18      if ((0x10000000 <= pa && pa + len <= 0x10000020) ||
19          (0x13000000 <= pa && pa + len <= 0x13004200) ||
20          (0x15000000 <= pa && pa + len <= 0x15000200)) {
21          memcpy((void *)va, (void *) (pa + KSEG1), len); //复制到va里
22          return 0;
23      }
24
25      return -E_INVALID;
26
27  }

```

## Exercise 5.2

在user/lib/syscall\_lib.c 中完成用户态的相应系统调用的接口。

```

1  int syscall_write_dev(void *va, u_int dev, u_int len) {
2      /* Exercise 5.2: Your code here. (1/2) */
3      return msyscall(SYS_write_dev, va, dev, len);
4  }
5
6  int syscall_read_dev(void *va, u_int dev, u_int len) {
7      /* Exercise 5.2: Your code here. (2/2) */
8      return msyscall(SYS_read_dev, va, dev, len);
9  }

```

## Exercise 5.3

Exercise 5.3 参考以上展示的内核态驱动，使用**系统调用**完成fs/ide.c 中的ide\_write函数，以及ide\_read 函数，实现对磁盘的读写操作。

**fs是用户态！ 要使用系统调用syscall**

```
1 // Overview:
2 // read data from IDE disk. First issue a read request through
3 // disk register and then copy data from disk buffer
4 // (512 bytes, a sector) to destination array.
5 //
6 // Parameters:
7 // diskno: disk number.
8 // secno: start sector number.
9 // dst: destination for data read from IDE disk.
10 // nsecs: the number of sectors to read.
11 //
12 // Post-Condition:
13 // Panic if any error occurs. (you may want to use 'panic_on')
14 //
15 // Hint: Use syscalls to access device registers and buffers.
16 // Hint: Use the physical address and offsets defined in
17 // 'include/drivers/dev_disk.h':
18 // 'DEV_DISK_ADDRESS', 'DEV_DISK_ID', 'DEV_DISK_OFFSET',
19 // 'DEV_DISK_OPERATION_READ',
20 // 'DEV_DISK_START_OPERATION', 'DEV_DISK_STATUS', 'DEV_DISK_BUFFER'
21 void ide_read(u_int diskno, u_int secno, void *dst, u_int nsecs) {
22     u_int begin = secno * BY2SECT;
23     u_int end = begin + nsecs * BY2SECT;
24
25     for (u_int off = 0; begin + off < end; off += BY2SECT) {
26         uint32_t temp = diskno;
27         /* Exercise 5.3: Your code here. (1/2) */
28         panic_on(syscall_write_dev(&temp, DEV_DISK_ADDRESS | DEV_DISK_ID, 4));
29         //设置下一次读写的磁盘编号(写入diskno)
30         uint32_t temp_off = begin + off; //写入off是begin + off
31         panic_on(syscall_write_dev(&temp_off, DEV_DISK_ADDRESS |
32 DEV_DISK_OFFSET, 4)); //设置下一次磁盘镜像偏移字节数(写入off)
33         u_int opt = DEV_DISK_OPERATION_READ; // opt=0
34         panic_on(syscall_write_dev(&opt, DEV_DISK_ADDRESS |
35 DEV_DISK_START_OPERATION, 4)); //写入0开始读磁盘
36
37         syscall_read_dev(&temp, DEV_DISK_ADDRESS | DEV_DISK_STATUS, 4); //读上一
38         次状态返回值(0是失败)
39         if(temp == 0) {
40             panic_on("fail to read.");
41         }
42         //读地址是dst+off
43         panic_on(syscall_read_dev(dst+off, DEV_DISK_ADDRESS | DEV_DISK_BUFFER,
44 BY2SECT)); //将对应off扇区的512bytes数据从设备缓冲区读入目标位置
45     }
46 }
```

```

1 // Overview:
2 // write data to IDE disk.
3 //
4 // Parameters:
5 // diskno: disk number.
6 // secno: start sector number.
7 // src: the source data to write into IDE disk.
8 // nsecs: the number of sectors to write.
9 //
10 // Post-Condition:
11 // Panic if any error occurs.
12 //
13 // Hint: Use syscalls to access device registers and buffers.
14 // Hint: Use the physical address and offsets defined in
15 // 'include/drivers/dev_disk.h':
16 // 'DEV_DISK_ADDRESS', 'DEV_DISK_ID', 'DEV_DISK_OFFSET', 'DEV_DISK_BUFFER',
17 // 'DEV_DISK_OPERATION_WRITE', 'DEV_DISK_START_OPERATION', 'DEV_DISK_STATUS'
18 void ide_write(u_int diskno, u_int secno, void *src, u_int nsecs) {
19     u_int begin = secno * BY2SECT;
20     u_int end = begin + nsecs * BY2SECT;
21
22     for (u_int off = 0; begin + off < end; off += BY2SECT) {
23         uint32_t temp = diskno;
24         /* Exercise 5.3: Your code here. (2/2) */
25         uint32_t temp_off = begin + off;
26         panic_on(syscall_write_dev(&temp, DEV_DISK_ADDRESS | DEV_DISK_ID, 4));
27         panic_on(syscall_write_dev(&temp_off, DEV_DISK_ADDRESS |
28 DEV_DISK_OFFSET, 4));
29         panic_on(syscall_write_dev(src+off, DEV_DISK_ADDRESS | DEV_DISK_BUFFER,
30 BY2SECT)); //先将对应扇区512bytes写入数据缓冲区
31         u_int opt = DEV_DISK_OPERATION_WRITE;
32         panic_on(syscall_write_dev(&opt, DEV_DISK_ADDRESS |
33 DEV_DISK_START_OPERATION, 4)); //再写入1启动写磁盘操作
34         syscall_read_dev(&temp, DEV_DISK_ADDRESS | DEV_DISK_STATUS, 4);
35         if(temp == 0) {
36             panic_on("fail to write.");
37         }
38     }
39 }

```

///// ↓ ↓ ↓ /////

## Exercise 5.4

文件系统需要负责维护磁盘块的申请和释放，在回收一个磁盘块时，需要更改位图中的标志位。如果要将一个磁盘块设置为free，只需要将位图中对应位的值设置为1即可。请完成fs/fs.c中的free\_block函数，实现这一功能。同时思考为什么参数blockno的值不能为0？

```

// Overview:
// Mark a block as free in the bitmap.
void free_block(u_int blockno) {
    // You can refer to the function 'block_is_free' above.
    // Step 1: If 'blockno' is invalid (0 or >= the number of blocks in 'super'), return.
    // Step 2: Set the flag bit of 'blockno' in 'bitmap'.
    // Hint: Use bit operations to update the bitmap, such as b[n / W] |= 1 << (n % W).
}

```

```

1 // Overview:
2 // Mark a block as free in the bitmap.
3 void free_block(u_int blockno) {
4     // You can refer to the function 'block_is_free' above.
5     // Step 1: If 'blockno' is invalid (0 or >= the number of blocks in
    'super'), return.
6     /* Exercise 5.4: Your code here. (1/2) */
7     if(blockno == 0 || blockno >= super->s_nblocks) { //super是超级块的指针，描述文件系统基本信息
8         return;
9     }
10    // Step 2: Set the flag bit of 'blockno' in 'bitmap'.
11    // Hint: Use bit operations to update the bitmap, such as b[n / w] |= 1 <<
    (n % w).
12    /* Exercise 5.4: Your code here. (2/2) */
13    bitmap[blockno / 32] |= (1 << (blockno % 32)); //bitmap数组每个元素含32bits,
    或上blockno对应32个里面的顺序
14 }

```

## Exercise 5.5

参照文件系统的设计，完成tools/fsformat.c中的create\_file函数，并阅读write\_directory函数（代码已在源文件中给出，不作为考查点），实现将一个文件或指定目录下的文件按照目录结构写入到target/fs.img的功能。

```

1 // Overview:
2 // Allocate an unused 'struct File' under the specified directory.
3 //
4 // Note that when we delete a file, we do not re-arrange all
5 // other 'File's, so we should reuse existing unused 'File's here.
6 //
7 // Post-Condition:
8 // Return a pointer to an unused 'struct File'.
9 // We assume that this function will never fail.
10 //
11 // Hint:
12 // Use 'make_link_block' to allocate a new block for the directory if there are
    no existing unused
13 // 'File's.
14 struct File *create_file(struct File *dirf) {

```

```

15     int nblk = dirf->f_size / BY2BLK;
16
17     // Step 1: Iterate through all existing blocks in the directory.
18     for (int i = 0; i < nblk; ++i) {
19         int bno; // the block number
20         // If the block number is in the range of direct pointers (NDIRECT), get
the 'bno'
21         // directly from 'f_direct'. Otherwise, access the indirect block on
'disk' and get
22         // the 'bno' at the index.
23         /* Exercise 5.5: Your code here. (1/3) */
24         if(i < NDIRECT) {
25             bno = dirf->f_direct[i];
26         } else {
27             bno = ((uint32_t*) (disk[dirf->f_indirect].data))[i];
28         }
29
30         // Get the directory block using the block number.
31         struct File *blk = (struct File *) (disk[bno].data);
32
33         // Iterate through all 'File's in the directory block.
34         for (struct File *f = blk; f < blk + FILE2BLK; ++f) {
35             // If the first byte of the file name is null, the 'File' is unused.
36             // Return a pointer to the unused 'File'.
37             /* Exercise 5.5: Your code here. (2/3) */
38             if(f->f_name == NULL) {
39                 return f;
40             }
41         }
42     }
43
44     // Step 2: If no unused file is found, allocate a new block using
'make_link_block' function
45     // and return a pointer to the new block on 'disk'.
46     /* Exercise 5.5: Your code here. (3/3) */
47     return (struct File*) (disk[make_link_block(dirf, nblk)].data);
48
49     return NULL;
50 }

```

## Exercise 5.6

fs/fs.c 中的 diskaddr 函数用来计算指定磁盘块对应的虚存地址。完成 diskaddr 函数，根据一个块的序号 (block number)，计算这一磁盘块对应的虚存的起始地址。（提示：fs/serv.h 中的宏 DISKMAP 和 DISKMAX 定义了磁盘映射虚存的地址空间）

```

1 // Overview:
2 // Return the virtual address of this disk block in cache.
3 // Hint: Use 'DISKMAP' and 'BY2BLK' to calculate the address.
4 void *diskaddr(u_int blockno) {
5     /* Exercise 5.6: Your code here. */
6     return DISKMAP + blockno*BY2BLK;
7 }

```

## Exercise 5.7

实现fs/fs.c 中的map\_block 函数，检查指定的磁盘块是否已经映射到内存，如果没有，分配一页内存来保存磁盘上的数据。相应地，完成unmap\_block 函数，用于解除磁盘块和物理内存之间的映射关系，回收内存。

（提示：注意磁盘虚拟内存地址空间和磁盘块之间的对应关系）。

```

1 // Overview:
2 // Allocate a page to cache the disk block.
3 int map_block(u_int blockno) {
4     // Step 1: If the block is already mapped in cache, return 0.
5     // Hint: Use 'block_is_mapped'.
6     /* Exercise 5.7: Your code here. (1/5) */
7     if(block_is_mapped(blockno)) {
8         return 0;
9     }
10
11     // Step 2: Alloc a page in permission 'PTE_D' via syscall.
12     // Hint: Use 'diskaddr' for the virtual address.
13     /* Exercise 5.7: Your code here. (2/5) */
14     syscall_mem_alloc(0, diskaddr(blockno), PTE_D); //使用系统调用分配内存，当前进程
是env; 将blockno转成对应虚拟地址
15 }

```

```

1 // Overview:
2 // Unmap a disk block in cache.
3 void unmap_block(u_int blockno) {
4     // Step 1: Get the mapped address of the cache page of this block using
'block_is_mapped'.
5     void *va;
6     /* Exercise 5.7: Your code here. (3/5) */
7     va = block_is_mapped(blockno);
8     if(va == NULL) {
9         return;
10    }
11    // Step 2: If this block is used (not free) and dirty in cache, write it
back to the disk
12    // first.

```

```

13 // Hint: Use 'block_is_free', 'block_is_dirty' to check, and 'write_block'
    to sync.
14 /* Exercise 5.7: Your code here. (4/5) */
15 if(!block_is_free(blockno) && block_is_dirty(blockno)) {
16     write_block(blockno);
17 }
18
19 // Step 3: Unmap the virtual address via syscall.
20 /* Exercise 5.7: Your code here. (5/5) */
21 syscall_mem_unmap(0, diskaddr(blockno)); //使用系统调用解除内存映射
22
23 user_assert(!block_is_mapped(blockno));
24 }

```

## Exercise 5.8

补全fs/fs.c 中的dir\_lookup 函数，查找某个目录下是否存在指定的文件。（使用file\_get\_block 函数）

```

1 // Overview:
2 // Find a file named 'name' in the directory 'dir'. If found, set *file to it.
3 //
4 // Post-Condition:
5 // Return 0 on success, and set the pointer to the target file in `*file`.
6 // Return the underlying error if an error occurs.
7 int dir_lookup(struct File *dir, char *name, struct File **file) {
8     int r;
9     // Step 1: Calculate the number of blocks in 'dir' via its size.
10    u_int nblock;
11    /* Exercise 5.8: Your code here. (1/3) */
12    nblock = dir->f_size / BY2BLK; //块数 = 文件总大小 ÷ 每块大小
13
14    // Step 2: Iterate through all blocks in the directory.
15    for (int i = 0; i < nblock; i++) {
16        // Read the i'th block of 'dir' and get its address in 'blk' using
        'file_get_block'.
17        void *blk;
18        /* Exercise 5.8: Your code here. (2/3) */
19        try(file_get_block(dir, i, &blk)); //将某个指定文件指向的磁盘块读入内存
20
21        struct File *files = (struct File *)blk;
22
23        // Find the target among all 'File's in this block.
24        for (struct File *f = files; f < files + FILE2BLK; ++f) {
25            // Compare the file name against 'name' using 'strcmp'.
26            // If we find the target file, set '*file' to it and set up its
        'f_dir'
27            // field.
28            /* Exercise 5.8: Your code here. (3/3) */
29            if(strcmp(f->f_name, name) == 0) {
30                *file = f;
31                f->f_dir = dir; //文件所属的文件目录

```



```

32         return 0;
33     }
34
35     }
36 }
37
38     return -E_NOT_FOUND;
39 }

```

## Exercise 5.9

请完成user/lib/file.c 中的open 函数。（提示：若成功打开文件，则该函数返回文件描述符的编号）

```

1  // Overview:
2  //  Open a file (or directory).
3  //
4  // Returns:
5  //  the file descriptor on success,
6  //  the underlying error on failure.
7  int open(const char *path, int mode) {
8      int r;
9
10     // Step 1: Alloc a new 'Fd' using 'fd_alloc' in fd.c.
11     // Hint: return the error code if failed.
12     struct Fd *fd;
13     /* Exercise 5.9: Your code here. (1/5) */
14     try(fd_alloc(&fd));    //先分配一个fd（文件描述符）
15
16     // Step 2: Prepare the 'fd' using 'fsipc_open' in fsipc.c.
17     /* Exercise 5.9: Your code here. (2/5) */
18     try(fsipc_open(path, mode, fd)); //将path对应文件的文件描述符共享到 fd 指针对应的
地址处。
19
20     // Step 3: Set 'va' to the address of the page where the 'fd''s data is
cached, using
21     // 'fd2data'. Set 'size' and 'fileid' correctly with the value in 'fd' as a
'Filefd'.
22     char *va;
23     struct Filefd *ffd;
24     u_int size, fileid;
25     /* Exercise 5.9: Your code here. (3/5) */
26     va = fd2data(fd);    //获得地址
27     ffd = (struct Filefd*)fd;    //强制转换
28     size = ffd->f_file.f_size;
29     fileid = ffd->f_fileid;
30
31     // Step 4: Alloc pages and map the file content using 'fsipc_map'.
32     for (int i = 0; i < size; i += BY2PG) {
33         /* Exercise 5.9: Your code here. (4/5) */
34         try(fsipc_map(fileid, i, va));    // i即偏移量offset
35

```

```

36     }
37
38     // Step 5: Return the number of file descriptor using 'fd2num'.
39     /* Exercise 5.9: Your code here. (5/5) */
40     return fd2num(fd);    //返回文件描述符fd的编号
41 }

```

## Exercise 5.10

参考user/lib/fd.c 中的write 函数, 完成read 函数。

```

1  // Overview:
2  //  Read at most 'n' bytes from 'fd' at the current seek position into 'buf'.
3  //
4  // Post-Condition:
5  //  Update seek position.
6  //  Return the number of bytes read successfully.
7  //  Return < 0 on error.
8  int read(int fdnum, void *buf, u_int n) {
9      int r;
10
11     // Similar to the 'write' function below.
12     // Step 1: Get 'fd' and 'dev' using 'fd_lookup' and 'dev_lookup'.
13     struct Dev *dev;
14     struct Fd *fd;
15     /* Exercise 5.10: Your code here. (1/4) */
16     if((r = fd_lookup(fdnum, &fd)) < 0 || (r = dev_lookup(fd->fd_dev_id, &dev))
17 < 0) {
18         return r;
19     }
20
21     // Step 2: Check the open mode in 'fd'.
22     // Return -E_INVAL if the file is opened for writing only (O_WRONLY).
23     /* Exercise 5.10: Your code here. (2/4) */
24     if((fd->fd_omode & O_ACCMODE) == O_WRONLY) {
25         return -E_INVAL;
26     }
27
28     // Step 3: Read from 'dev' into 'buf' at the seek position (offset in 'fd').
29     /* Exercise 5.10: Your code here. (3/4) */
30     r = dev->dev_read(fd, buf, n, fd->fd_offset);    //Dev中读取(read) int
31     (*dev_read)(struct Fd *, void *, u_int, u_int);
32
33     // Step 4: Update the offset in 'fd' if the read is successful.
34     /* Hint: DO NOT add a null terminator to the end of the buffer!
35      * A character buffer is not a C string. Only the memory within [buf,
36      buf+n) is safe to
37      * use. */
38     /* Exercise 5.10: Your code here. (4/4) */
39     if(r > 0) {    // read成功则更新offset
40         fd->fd_offset += r;

```

```

38     }
39
40     return r;
41 }

```

## Exercise 5.11

完成fs/serv.c 中的serve\_remove 函数。

```

1  // Overview:
2  //  Serve to remove a file specified by the path in `req`.
3  void serve_remove(u_int envid, struct Fsreq_remove *rq) {
4      // Step 1: Remove the file specified in 'rq' using 'file_remove' and store
      // its return value.
5      int r;
6      /* Exercise 5.11: Your code here. (1/2) */
7      r = file_remove(rq->req_path);
8
9      // Step 2: Respond the return value to the requester 'envid' using
      // 'ipc_send'.
10     /* Exercise 5.11: Your code here. (2/2) */
11     ipc_send(envid, r, 0, 0);
12 }

```

## Exercise 5.12

完成user/lib/fsipc.c 中的fsipc\_remove 函数。

```

1  // Overview:
2  //  Ask the file server to delete a file, given its path.
3  int fsipc_remove(const char *path) {
4      // Step 1: Check the length of 'path' using 'strlen'.
5      // If the length of path is 0 or larger than 'MAXPATHLEN', return -
      // E_BAD_PATH.
6      /* Exercise 5.12: Your code here. (1/3) */
7      if(!(*path) || strlen(path) > MAXPATHLEN) {
8          return -E_BAD_PATH;
9      }
10
11     // Step 2: Use 'fsipcbuf' as a 'struct Fsreq_remove'.
12     struct Fsreq_remove *req = (struct Fsreq_remove *)fsipcbuf;
13
14     // Step 3: Copy 'path' into the path in 'req' using 'strcpy'.
15     /* Exercise 5.12: Your code here. (2/3) */
16     strcpy(req->req_path, path);    //strcpy(dst, src);
17
18     // Step 4: Send request to the server using 'fsipc'.
19     /* Exercise 5.12: Your code here. (3/3) */
20     return fsipc(FSREQ_REMOVE, req, 0, 0);
21 }

```

## Exercise 5.13

---

完成user/lib/file.c 中的remove 函数。

```
1 // Overview:
2 // Delete a file or directory.
3 int remove(const char *path) {
4     // Your code here.
5     // Call fsipc_remove.
6
7     /* Exercise 5.13: Your code here. */
8     return fsipc_remove(path);
9 }
```