

עבודת סיום קורס מערכות הפעלה 2022

סימולציה של Disk File System - 'מערכת לניהול הדיסק'

תאריך הגשה 4.8.2022

תזכורת - עבודה עצמאית, מצאתי פתרון ברשת או במקום אחר או אצל סטודנט אחר או שנה אחרת – זה לא עבודה עצמאית, וגם לא משהו דומה... אנחנו גם נבדוק את זה. (בבקשה, בבקשה, בבקשה!) זה רק בשבילכם, לא תעבדו עצמאית לא תלמדו... אין קיצורי דרך)

תיאור המערכת

סימולציה זוהי סביבת הדמיה בתוכנה לאירועים ופעולות הקורים במערכת אמיתית (חומרה או תוכנה).

מבוא ורקע :

מערכת ניהול הדיסק במערכת ההפעלה (Disk File System) היא הדרך שבה שמות קבצים מיקומם ותוכנם מאורגנים על גבי הדיסק הקשיח. ללא מערכת הקבצים, המידע מאוחסן לא יהיה מאורגן לקבצים בודדים ויהיה בלתי אפשרי לאתר ולגשת לתוכנם.

המשתמש "הפשוט" אשר משתמש בתוכנת הוורד למשל, רואה לפניו קובץ וורד כאשר מיקום הקובץ הפיזי על-גבי הדיסק, אינו מעניינו של המשמש. וזו בדיוק תפקידה של 'מערכת לניהול הדיסק' במערכת ההפעלה, למפות את כל חלקי הקובץ אשר מאוחסנים על-גבי הדיסק. חלקי הקובץ שמורים ביחידות קטנות אשר נקראות 'בלוקים' ומאורגנים לכדי יחידה לוגית אחת. היא הקובץ. חלקי הקובץ אינם נמצאים באופן ישיר ורציף על הדיסק אלא מפוזרים על פני הדיסק. המיפוי של חלקי הבלוקים האלה לכדי קובץ שלם על ידי מערכת הקבצים בדומה לתרגיל מספר 6 אשר ביצעתם בסוף הסמסטר -- בתרגיל זה אנו נדרשים לסמלץ מערכת קבצים במערכת מחשב קטנה עם דיסק קטן ותיקיה בודדת ונממש את כל הפעולות אשר מערכת הפעלה עושה על הדיסק

נקודות מרכזיות בתרגיל:

- הדיסק שלנו יהיה למעשה קובץ! (בדומה למרחב השחלוף של הזיכרון בתרגיל 5). בגודל של 256 תווים בשם **"DISK_SIM_FILE.txt"**
- מערכת קבצים זו תכיל רק תיקייה אחת וכל הקבצים ייווצרו תחת תיקייה זו. (לא נממש תת תיקיות)
- מערכת ניהול הדיסק אותה נסמלץ היא **Indexed allocation**. בקצרה: לכול קובץ יהיה בלוק שמכיל מצעים לבלוקים של נתונים (ניתן דוגמא, וניתן לחזור על החומר בשיעור)
- בתרגיל חובה לממש שלוש מחלקות/מבני נתונים **fsDisk**, **FileDescriptor**, **FsFile** כמו כן, נגדיר 8 פונקציות חובה למימוש במחלקה **fsDisk**.
 - את 8 הפעולות הבאות: **FsFormat**, **CreateFile**, **OpenFile**, **CloseFile**, **DeleteFile**, **WriteToFile**, **ReadFromFile**, **listAll**.
- ה-**main** של התרגיל נתון לכם וכן קוד ראשוני של המחלקות, וכן הפלט הנדרש..

בתרגיל זה אנו נממש יש לנהל שלושה מבני נתונים (מחלקות):

1. מידע על קובץ בודד ברמה "הנמוכה" ישמר במבנה הנתונים -- **FsFile**
- תפקידו לשמור מידע בסיסי על הקובץ (גודל קובץ, באיזה בלוקים נישמר הקובץ.. ע"י מצביע ל indexBlock וכו')
2. קישור בין שם הקובץ לFsFile שלו, ישמר ב **FileDescriptor**
- מבנה הנתונים הפשוט הנ"ל שומר את שם קובץ ומצביע ל-FsFile של הקובץ.
(תיקיה במערכת היא למעשה מערך של FileDescriptor)
3. הדיסק עצמו ופעולות שניתן לעשות על הדיסק ישמור ב **fsDisk**
- הדיסק עצמו, שומר את כל נתוני הדיסק.

הרחבה על כל אחד ממבנה הנתונים/מחלקות

כעת, נסביר בהרחבה על כל אחד ממבני הנתונים, אחר-כך נסביר על על אחת מ-8 הפעולות וכן נתאר את הממשק

1. fsFile

כאמור ה-fsFile הוא מבנה נתונים אשר שומר מידע בסיסי ברמה "הנמוכה" על הקובץ בשיטת index allocation

```
class FsFile {
    int file_size;
    int block_in_use;
    int index_block;
    int block_size;

public:
    FsFile(int _block_size) {
        file_size = 0;
        block_in_use = 0;
        block_size = _block_size;
        index_block = -1;
    }

    int getfile_size(){
        return file_size;
    }
}
```

במחלקה fsFile אתם נדרשים להגדיר שלוש שדות חובה: file_size, index_block, block_size - נסביר אותם:

- block_size - מספר התווים שאפשר לשמור בכל בלוק
 - index_block - שדה זה שומר את מיקום ה index_block - כלומר את מספר הבלוק אשר מחזיק את המצביעים לבלוקים של נתוני הקובץ.
 - file_size - גודל הקובץ (בתווים)
- אופציונאלי - block_in_use - מספר הבלוקים בשימוש

ה-constructor מאתחל את גודל הקובץ (fileSize) ומספר הבלוקים לשימוש (block_in_used) לאפס. כמו כן מקצה ומאתחל מספר ה index_block ל -1. (כלומר, טרם הוקצה).

מעבר לזה אתם רשאים/נדרשים להוסיף עוד פונקציות למחלקה זו.

מספר הבלוקים המקסימלי של קובץ במערכת שלנו יוכל להיות הם:

block_size

מעבר לזה - במערכת שלנו, לא נוכל לשמור יותר בלוקים עבור קובץ בודד!

ובהתאם, גודל קובץ (מספר התווים אשר ניתן לשמור בקובץ יהיה

`block_size * block_size`

2. `FileDescriptor`

המחלקה בפשטות שומרת מידע מצומד של שם קובץ ומצביע ל-`FsFile` של הקובץ. בנוסף, שומרת המחלקה משתנה בוליאני `inUse` שערכו שווה ל `true` כאשר פותחים את הקובץ ושווה ל-`false` כאשר סוגרים אותו. (שימו לב, סגירת קובץ זה לא מחיקת קובץ).

נשים לב שהתיקייה (`MainDir`) במערכת היא למעשה מערך של `FileDescriptor`ים. במערכת שלנו תהיה תיקייה יחידה, [עוד בנושא זה - במבנה הנתונים הבא....]

חתימה של תחילת המחלקה וה `constructor`:

```
// =====
class FileDescriptor {
    string file_name;
    FsFile* fs_file;
    bool inUse;

public:
    FileDescriptor(string FileName, FsFile* fsi) {
        file_name = FileName;
        fs_file = fsi;
        inUse = true;
    }

    string getFileName() {
        return file_name;
    }
}
```

אין לשנות מבנה נתונים זה.

3. fsDisk

```
class fsDisk {
    FILE *sim_disk_fd;
    bool is_formated;

    // BitVector - "bit" (int) vector, indicate which block in the disk is free
    // or not. (i.e. if BitVector[0] == 1, means that the
    // first block is occupied.
    int BitVectorSize;
    int *BitVector;

    // (5) MainDir --
    // Structure that links the file name to its FsFile

    // (6) OpenFileDescriptors --
    // when you open a file,
    // the operating system creates an entry to represent that file
    // This entry number is the file descriptor.
}
```

זאת המנה העיקרית.... מבנה הנתונים האחרון שחובה להגדיר בתרגיל, הוא הדיסק עצמו. מחלקה בשם fsDisk.

במחלקה זו, ישנם 6 שדות חובה הבאים: sim_disk_fd, is_formated, BitVectorSize, BitVector, MainDir, OpenFileDescriptors

כאשר הטיפוס 4 שדות נתון לכם ו 2 נוספים (MainDir, OpenFileDescriptors) לבחירתכם, בהתאם למימוש שתבחרו.

- המשתנה הראשון sim_disk_fd הוא מצביע על הקובץ מסוג של FILE שאותו נפתח בעזרת פקודה fopen (ראו constructor) זה *הדיסק שישמש אותנו לסימולציה.
 - המשתנה הבא, הוא is_formated - משתנה בוליאני, אשר מציין האם הדיסק הוא כבר עבור פורמט או לא (מדליקים אותו ל-true בסיום הפונקציה fsFormat (ראו בהמשך).
 - משתנה נוסף הוא מערך בשם BitVector מסוג int, כל תא i במערך מציין האם הבלוק מספר i תפוס/בשימוש, כן או לא.
 - משתנה נוסף הוא MainDir, זו "מערך" אשר מקשרת בין שם הקובץ ל-FsFile שלו. כאשר המימוש הספציפי נתון לשיקול דעתכם.
 - כמו כן, יהיה לנו וקטור בשם OpenFileDescriptors, זה למעשה זה מערך של FileDescriptor שהם כל הקבצים הפתוחים, כל הקבצים שפתחנו בסימולציה. גם כאן לא הגדנו לכם באיזה מבנה נתונים להשתמש וזה לשיקול דעתכם.
- שאלה: למה בכלל מעניין אותנו מבנה נתונים זה והאם אין כאן כפילות של MainDir ? תשובה: בעזרת מבנה זה נוכל לדעת איזה file descriptors פתוחים \ אילו סגורים ובמקרה של פתיחת קובץ חדש נוכל לדעת איזה מספר file descriptors חדש יש להקצות. כאשר MainDir מחזיק את כל הקבצים גם הסגורים...

```
// -----
fsDisk() {
    sim_disk_fd = fopen(DISK_SIM_FILE , "r+");
    assert(sim_disk_fd);

    for (int i=0; i < DISK_SIZE ; i++) {
        int ret_val = fseek ( sim_disk_fd , i , SEEK_SET );
        ret_val = fwrite( "\0" , 1 , 1, sim_disk_fd);
        assert(ret_val == 1);
    }

    fflush(sim_disk_fd);
    is_formated = false;
}
```

כמו כן נתנו לכם בסיס לקונסטרקטור שבו יש להשתמש - כל תפקידו בשלב זה הוא לפתוח את קובץ הסימולציה של הדיסק ולאפס את תוכנו ואת שאר המשתנים..

אתם רשאים להוסיף עוד שדות ופונקציות לפי הבנתכם.

כעת, סיימנו להגדיר את מבני הנתונים העיקריים, נפנה להגדרת ה-main והפונקציות שחובה שיש לממש

```
int main() {
    int blockSize;
    string fileName;
    char str_to_write[DISK_SIZE];
    char str_to_read[DISK_SIZE];
    int size_to_read;
    int _fd;
    int wrote;

    fsDisk *fs = new fsDisk();
    int cmd;
    while(1) {
        cin >> cmd;
        switch (cmd)
        {
            case 0: // exit
                delete fs;
                exit(0);
                break;

            case 1: // list-file
                fs->listAll();
                break;

            case 2: // format
                cin >> blockSize;
                fs->fsFormat(blockSize);
                break;

            case 3: // creat-file
                cin >> fileName;
                _fd = fs->CreateFile(fileName);
                cout << "CreateFile: " << fileName << " with File Descriptor #: " << _fd << endl;
                break;

            case 4: // open-file
                cin >> fileName;
                _fd = fs->OpenFile(fileName);
                cout << "OpenFile: " << fileName << " with File Descriptor #: " << _fd << endl;
                break;

            case 5: // close-file
                cin >> _fd;
                fileName = fs->CloseFile(_fd);
                cout << "CloseFile: " << fileName << " with File Descriptor #: " << _fd << endl;
                break;

            case 6: // write-file
                cin >> _fd;
                cin >> str_to_write;
                wrote = fs->WriteToFile(_fd, str_to_write, strlen(str_to_write));
                cout << "Wrote: " << wrote << " Char's into File Descriptor #: " << _fd << endl;
                break;

            case 7: // read-file
                cin >> _fd;
                cin >> size_to_read;
                fs->ReadFromFile(_fd, str_to_read, size_to_read);
                cout << "ReadFromFile: " << str_to_read << endl;
                break;

            case 8: // delete file
                cin >> fileName;
                _fd = fs->DelFile(fileName);
                cout << "DeletedFile: " << fileName << " with File Descriptor #: " << _fd << endl;
                break;

            default:
                break;
        }
    }
}
```

בתרגיל זה, ממשק ה-main נתון ובנוי על לולאה אשר כל פעם קולטת פקודה (מספר) מהמשתמש שהוא מספר בין אפס לשמונה.

- אופציה מספר אפס: מחיקת כל הדיסק ויציאה.
- אופציה מספר אחת: יש להדפיס את כל הקבצים שקיימים בדיסק void listAll(). הפונקציה הזאת נתונה לכם. הפונקציה מדפיסה את רשימת הקבצים שנוצרו בדיסק וכן את תוכן הדיסק¹.
- אופציה מספר שתיים: פירמוט הדיסק - זימון הפונקציה fsFormat לפרמוט הדיסק. לצורך זה יש לקלוט מהמשתמש את גודל הבלוק בדיסק.
- אופציה מספר שלוש: יצירת קובץ ופתיחת קובץ - זימון הפונקציה createFile, לצורך זה יש לקלוט מהמשתמש "שם קובץ", הפונקציה CreateFile מייצרת קובץ חדש במערכת. (רמז למימוש: פונקציה זו תהיה אחראית ליצירת fsFile וכן לעדכן את מבני הנתונים MainDir וOpenFileDescriptors). הפונקציה תחזיר את file descriptor של הקובץ שנפתח (רמז: זה למעשה מיקומו בווקטור OpenFileDescriptors). הפונקציה גם תבדוק שהדיסק כבר אותחל ואם לא - תחזיר 1- (מינוס 1)
- אופציות מספר ארבע וחמש הן פתיחה וסגירה של קובץ. אופציה מספר ארבע: פתיחת קובץ OpenFile מחזירה את ה file descriptor, יש לוודא שהקובץ קיים ולא פתוח כבר... (כאמור אופציה מספר שלוש יוצרת את הקובץ וגם פותחת אותו). אופציה מספר חמש נותנת לנו אפשרות לסגור את הקובץ CloseFile בהינתן file descriptor. כמובן שהפונקציה צריכה לוודא שיש קובץ כ"ל ושהוא פתוח. בכל מקרה של שגיאה הפונקציה תחזיר 1- כ- string כאשר CloseFile או OpenFile into
- אופציות מספר שש ושבע הן, כתיבה וקריאה מקובץ: כאשר אנחנו רוצים לכתוב לקובץ ראשית לקלוט מהמשתמש מספר file descriptor של קובץ שאליו יש לכתוב, וסטרינג שאותו רוצים לרשום לתוך הקובץ. בהינתן שני אלו, מזמנים את הפונקציה WriteToFile שתכתוב נתונים לקובץ. חלק מהבדיקות שכמובן יש לוודא בפונקציה זו הם: (1) שיש מספיק מקום בדיסק, (2) בקובץ, (3) שהקובץ נפתח (4) ושהדיסק אותחל ואם לא תחזיר 1- (רמז למימוש WriteToFile: הפונקציה צריכה למצוא בלוקים פנויים בדיסק כדי לרשום לתוכם את הנתונים. אם כבר כתבו לקובץ זה בעבר, אולי נשאר מקום בבלוק האחרון שכבר הוקצה לקובץ זה ואז יש להשתמש בו... ובכל מקרה - בכל פעם יש להקצות מספר מינימלי של בלוקים הדרושים כדי לספק את הכתיבה הדרושה) אופציה מספר שבע לקריאה מקובץ ReadFromFile - ראשית שוב יש לקלוט מהמשתמש file descriptor של קובץ שממנו יש לקרוא וכן את מספר התווים שיש לקרוא. הפונקציה תיגש לקובץ המתאים, משם לבלוקים המתאימים (את הבלוקים נמצא דרך מבנה הנתונים FsFile של הקובץ) הפונקציה תחזיר את התווים שביקשנו ותדפיס אותם (כבר ממומש לכם ההדפסה: -) ...דברים נוספים שהפונקציה צריכה לבדוק? (תחשבו).
- אופציה מספר שמונה היא מחיקת קובץ. תקבל את שם הקובץ ותמחק את כל הנתונים שקשורים אליו. יש למחוק גם את ה-File שלו מתוך המאגר.

¹ תוכלו להעזר בפונקציה זו כדי ללמוד גם איך לגשת לדיסק לסרוק אותו ולקרוא ממנו נתונים,

שימו לב שפונקציית ההדפסה נתונה לכם (וזאת כדאי שההדפסה תצאה אחידה לכולם בבדיקה), אך יחד עם זאת יש כמה שורות קוד להשלים... וזאת בהתאם למימוש שלכם לשמירת ה file descriptors ... הדפיסו את השם הקובץ והאם הוא בשימוש כן \ לא .

```
// -----
fsDisk() {
    sim_disk_fd = fopen(DISK_SIM_FILE , "r+");
    assert(sim_disk_fd);

    for (int i=0; i < DISK_SIZE ; i++) {
        int ret_val = fseek ( sim_disk_fd , i , SEEK_SET );
        ret_val = fwrite( "\0" , 1 , 1, sim_disk_fd);
        assert(ret_val == 1);
    }

    fflush(sim_disk_fd);
    is_formated = false;
}
```

פונקציה מאוד חשובה decToBinary - מתי פונקציה זו שימושית? כאשר רוצים לשמור את מספרי הבלוקים של ה- לדיסק. יש להמיר את מספר הבלוק שמור במשתנה n לצורתו הבינארית כ-char שיישמר בתו c.

```
// =====
void decToBinary(int n, char &c)
{
    // array to store binary number
    int binaryNum[8];

    // counter for binary array
    int i = 0;
    while (n > 0)
    {
        // storing remainder in binary array
        binaryNum[i] = n % 2;
        n = n / 2;
        i++;
    }

    // printing binary array in reverse order
    for (int j = i - 1; j >= 0; j--)
    {
        if (binaryNum[j] == 1)
            c = c | 1u << j;
    }
}
```

זהו ?

הגשה כרגיל... כולל README

בהצלחה!