Operating Systems – 234123 Homework Exercise 3 – Dry

Submitters:

Tom Guy

315155671

tom.guy@campus.technion.ac.il

Nikita Ivanov

322245101

Nikitaiv@campus.technion.ac.il

שאלה 1 - זיכרון:

עדן, זמרת פופולרית, סבלה ממחסור בזיכרון פיזי במחשב שלה (בעל מעבד IA-32 וזיכרון פיזי בגודל 4GB עדן, זמרת פופולרית, סבלה ממחסור בזיכרון פיזי במחשב שלה (בעל מעבד המרחיב את מרחב הזיכרון הפיזי מ-32 ל-40 ביט. כתוצאה מכך, במימוש של עדן יש שלוש רמות תרגום בטבלת הדפים.

שאר נתוני המעבד של עדן זהים לאלו של מעבד IA-32, כלומר נתוני המערכת החדשה הם:

רוחב כתובת וירטואלית	32bit
רוחב כתובת פיזית	40bit
גודל דף/מסגרת/מגירה	4KB
גודל מסגרת של טבלת דפים (בכל הרמות)	4KB
מספר ביטים לדגלים והרשאות בכל כניסה בטבלת הדפים	12bit

סעיף 1

בהנחה שגודל כניסה בטבלת הדפים מעוגל למעלה לחזקה שלמה של 2, מהו אופן חלוקת הכתובת הוירטואלית לשדות בהנחה שגודל כניסה בטבלת הדפים מעוגל למעלה לחזקה שלמה של 2, מהו אופן (page walk) נמקי.

index3	index2	index1	offset	
2	9	9	12	<mark>8.</mark>
2	10	10	10	ם.
1	9	10	12	٦.
4	9	9	10	7.
2	10	10	12	ה.
2	9	9	10	٦.

נימוק:

תשובה:

ראשית, מכך שגודל הדף נשאר זהה – 4KB, גם החישוב של מציאת offset בתוך הדף יישאר זהה ולכן דרושים 12 ביטים למציאת מיקום הכתובת הפיזית המבוקשת בתוך הדף. כעת נבין את ההיררכיה הרצויה על מנת לתמוך בכתובות פיזיות בגודל 40 ביטים. גודל כל מוקום הכתובת מעוגל לחזקה שלמה של 2 ולכן יהיה 64 ביטים. גודל כל page table הוא כגודל דף – 4KB ומכך שכל כתובת בו היא כל כתובת מעוגל לחזקה שלמה של 2 ולכן יהיה 64 ביטים. ציטים כדי למפות כל PTE ביטים פון יש לנו מספיק ביטים כדי למפות כל page table יש לנו מספיק ביטים כדי לעשות 2 page table כאלה ול-page table האחרון יישארו 2 ביטים.

2 סעיף

לבעלה של עדן, שוקי, אין שום תואר מהטכניון, ולמרות זאת הוא הבחין כי המימוש של עדן בזבזני בגלל שגודל הכניסות בטבלת הדפים מעוגל למעלה לחזקה שלמה של 2.

מהו הגודל המינימלי האפשרי של כניסה בטבלת הדפים אם לא מעגלים אותו למעלה? נמקי.

- 3 (a
- 4 (b
- 5 (c
- 6 (d
- 7 (e

אף תשובה לא נכונה (f

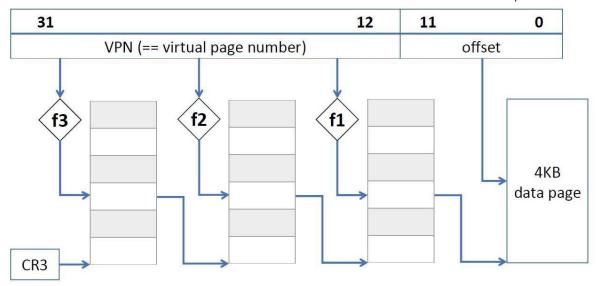
<u>נימוק</u>:

ביטים למיפוי (offset לכן כל כניסה בטבלת הדפים שמכילה 12 ביטים של דגלים ו- 28 (± 40 -12, 12 ביטים של ביטים למיפוי), ביטים למיפוי כתובת יכולה להישמר בלפחות חמישה בייטים.

<u>3 סעיף</u>

בהמשך לסעיף הקודם, שוקי (בעלה של עדן) הציע מימוש חדש לטבלת הדפים שבו כל כניסה בטבלת הדפים (בכל הרמות) היא בגודל המינימלי מהסעיף הקודם. במימוש של שוקי, כמו במימוש המקורי של טבלת הדפים במעבדי אינטל, דפים סמוכים בזיכרון הווירטואלי נשמרים בכניסות סמוכות בטבלת הדפים. שוקי הבחין שבמימוש החדש הכתובת הווירטואלית אינה מתפרקת לשדות של אינדקסים ויש צורך בחישובים מורכבים על מנת למצוא את האינדקס המתאים בכל טבלה (כלומר בכל רמה בעץ). להלן שרטוט

הממחיש את אופן התרגום:



בשרטוט רואים שלוש פונקציות (f1, f2, f3) המקבלות את מספר הדף הווירטואלי VPN בשרטוט רואים שלושה הרמות בטבלת הדפים.

בכל הסעיפים הבאים, הפעולות חלוקה "\" ומודולו "%" הן פעולות בשלמים. למשל:

מהי הפונקציה f1? נמקי.

- $f_1(vpn) = vpn/819$.a
- $f_1(vpn) = vpn\%819$.b
- $f_1(vpn) = (vpn/819)\%819$.c
- $f_1(vpn) = (vpn/819)/819$.d
- $f_1(vpn) = (vpn\%819)/819$.e
- $f_1(vpn) = ((vpn/819)\%819)/819$.f

תשובה: לסעיף הנוכחי ולשניים אחריו

מסעיף קודם, ראינו כי גודלה של כל כתובת (כניסה בטבלה) הוא 5 בייטים ולכן יש סה"כ 819.2 = 4096/5 כתובות אפשריות בכל דף – כמובן שנשתמש במספר שלם של כתובות ולכן יש סה"כ 819 כתובות בדף. את המיפוי עבור כתובת וירטואלית כללית נבצע בצורה הבאה:

 $0 \le a, b, c \le 818$ באשר $a * 819^2 + b * 819 + c$ נציג את הכתובת בצורה הבאה:

- <u>המיפוי לפונקציות יהיה:</u>
- f1(vpn) = c = vpn%819 \circ
- f2(vpn) = b = (vpn/819)%819 \circ
- f3(vpn) = a = (vpn/819)/819 \circ

סעיף <u>4</u> מהי הפונקציה

?f2 <u>נמקי</u>.

- $f_2(vpn) = vpn/819$.a
- $f_2(vpn) = vpn\%819$.b
- $f_2(vpn) = (vpn/819)\%819$.c
- $f_2(vpn) = (vpn/819)/819$.d
- $f_2(vpn) = (vpn\%819)/819$.e
- $f_2(vpn) = ((vpn/819)\%819)/819$.f

תשובה:

ענינו גם על סעיף זה מעלה.

<u>. מהי הפונקציה 51</u> מהי הפונקציה

- $f_3(vpn) = vpn/819$.a
- $f_3(vpn) = vpn\%819$.b
- $f_3(vpn) = (vpn/819)\%819$.c
- $f_3(vpn) = (vpn/819)/819$.d
- $f_3(vpn) = (vpn\%819)/819$.e
- $f_3(vpn) = ((vpn/819)\%819)/819$.f

תשובה:

ענינו גם על סעיף זה מעלה.

<u>6 סעיף</u>

מה היתרון של המערכת שהציע שוקי על פני המערכת שהציעה עדן? נמקי.

- ותר. של מרחב זיכרון וירטואלי גדול יותר. (a
 - של מיפוי של מרחב זיכרון פיזי גדול יותר. (b
- ר-TLB אפקטיבי יותר בגלל שהוא מכסה יותר זיכרון.
- d) טבלאות הדפים של תהליכי משתמש תופסות נפח קטן יותר בזיכרון.
 - e) פחות פרגמנטציה חיצונית, כלומר יותר זיכרון רציף.
 - f) אף תשובה אינה נכונה.

<u>נימוק</u>:

במימוש של עדן, כל טבלת דפים הכילה PTEs' כל היותר (עבור $\frac{4096}{8} = 512$ במימוש של שוקי כל טבלה מכילה PTEs' פמימוש אוקי כל טבלה מכילה PTEs' 819.

7 סעיף

לכל אחד מהסעיפים הבאים, סמני האם התשובה נכונה או לא, **ונמקי**.

<u>הגדרה:</u> מצביע **חוקי** הינו כזה שמצביע על כתובת בזיכרון שנמצאת בתוך אזור המוקצה בשביל התהליך\חוט גרעין\.. שמחזיק את המצביע.

1) אם למחשב יש 16GB של **RAM**, יתכן שלתוכנית <u>משתמש</u> כלשהי יש מצביע חוקי לכתובת הגדולה מ- (HEX בסיס 0x20000001).

? רמז: האם הכתובת ששמורה בתוך מצביע היא פיזית או ווירטואלית

<u>תשובה</u>: <mark>נכון</mark> \ לא נכון.

נימוק:

הכתובת היא מצביע בתוכנית - כתובת וירטואלית ולכן יתכן בהחלט כי המשתמש יקבל אותה בתור מצביע בעוד שהיא לא קיימת בזיכרון הפיזי, אד היא תמופה לכתובת פיזית הקטנה מ0x20000001.

2) אם למחשב יש 16GB של זיכרון **ווירטואלי**, יתכן שיש מצביע חוקי (לא בהכרח לתהליך משתמש) לכתובת הגדולה מ 16,000,000,000 (בסיס דצימלי).

תשובה: נכון \ לא נכון.

נימוק:

ייתכן בהחלט. לדוגמה, אם גודל כתובת פיזית גדול מגודל כתובת וירטואלית (יש יותר זיכרון פיזי מוירטואלי), כמו לדוגמה בסעיף א' (רק עם גודל כתובת שונה) יתכן כי הכתובות הגדולות מ-16,000,000,000 קיימות ותהליכי גרעין שנגישים לכתובות בצורה ישירה ללא זיכרון וירטואלי ייגשו אליהן בצורה חוקית כחלק מריצת תהליך.

0ממפות מחשב גודל דף = גודל מסגרת = 2KB, יתכן שהכתובת הוויראלית 00x800 ו 00x808 ממפות כתובת במרחק גדול מ00x80 אחת מהשנייה בזיכרון הפיזי.

תשובה: נכון \ לא נכון.

נימוק:

 $\underline{11}$ חייב להיות $\underline{01}$ ולכן גודל ה- $\underline{01}$ ולכן ביטים.

עבור 0x800 ועבור 0x808 כל הביטים אחרי הביט ה-11 שווים ורק ה-offset ביניהם שונה. לכן הם ממופים לאותו offset בגלל שההפרש ב-offset הוא 8, ההפרש בין שתי הכתובות הפיזיות הוא בדיוק 8.

4) מנגנון הזיכרון הווירטואלי (כמו שהוא) מאפשר לנו לדמות דיסק (אחסון עמיד) גדול יותר ממה שיש לנו בפועל.

<u>תשובה</u>: נכון \ <mark>לא נכון.</mark>

<u>נימוק</u>:

נשים לב כי אמנם חלק מהמטרה של זיכרון וירטואלי הוא לדמות שמירת זיכרון רב יותר ממה שיש למחשב לתת, אבל זה לא אפשרי לאחסן יותר מידע מכמות הזיכרון הפיזי שיש לנו ולכן לא יווצר מצב ששמור לנו יותר מידע בזכרון הוירטואלי מאשר כמות הזיכרון הפיזי (ושמירת חלק מהמידע על ה-RAM לא נחשב כדימוי לדיסק גדול יותר).

: שאלה 2 – ניהול זיכרון

שאלה זו כתובה באנגלית מאחר ושהיא מכילה לא מעט מושגים ושמות אשר קל יותר לבטאם באנגלית. נא לפתור אותה באיזה שפה שתרצנה\ו.

In the wet part of this homework, you implemented an interface that manages dynamic memory in for a process.

In this part of the homework, you will analyze the existing malloc() (from <cstdlib>) while learning about some new Linux tools.

NOTE: Do NOT submit code you write in this homework with your wet submission. Simply copy your code to your dry submission file, wherever requested.

Section 1:

1. Look up the "strace" utility online, read a little bit, and try to use it yourself by running `strace ls` in your OS terminal. Finally, explain here in a few words what it does.

The "strace" utility collects all the system calls the specified command used while running, and returns them out after the command exists.

2. Write a simple program in C that receives a number "x" from the command line and allocates (using malloc()) a block of memory that is "x" bytes long. You can assume there's always one input it will always be a positive integer. Run strace with your compiled program.

Finally, attach the code of the program and a screenshot of the output of running strace with your compiled program below

```
#include <stdio.h>
#include <stdib.h>

int main(int argc, char* argv[]) {
    if (argc != 2)
    {
       printf("Usage: %s <number> :(\n", argv[0]);
    }
    int x = atoi(argv[1]);
    malloc(x);

return 0;
}
```

```
~/Desktop$ strace ./a.out 8
.out", ["./a.out", "8"], 0x7ffd3f789418 /* 46 vars */) = 0
             ./a.out", ["./a.out",
brk(NULL)
                                                                  = 0x55bbeb485000
access("/etc/ld.so.nohwcap", F_OK) = -1 ENOENT (No such file access("/etc/ld.so.preload", R_OK) = -1 ENOENT (No such file openat(AT_FDCWD, "/etc/ld.so.cache", O_RDONLY|O_CLOEXEC) = 3 fstat(3, {st_mode=S_IFREG|0644, st_size=96204, ...}) = 0 mmap(NULL, 96204, PROT_READ, MAP_PRIVATE, 3, 0) = 0x7f1287656000
                                                                = -1 ENOENT (No such file or directory)
= -1 ENOENT (No such file or directory)
access("/etc/ld.so.nohwcap", F_OK) = -1 ENOENT (No such file or direct
openat(AT_FDCWD, "/lib/x86_64-linux-gnu/libc.so.6", O_RDONLY|O_CLOEXEC) = 3
                                                                  = -1 ENOENT (No such file or directory)
 ead(3, "\177ELF\2\1\1\3\0\0\0\0\0\0\0\3\0>\0\1\0\0\240\35\2\0\0\0\0\0\0"..., 832) = 832
 fstat(3, {st mode=S_IFREG|0755, st size=2030928, ...}) = 0
nmap(NULL, 8192, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_ANONYMOUS, -1, 0) = 0x7f1287654000
 map(NULL, 4131552, PROT_READ|PROT_EXEC, MAP_PRIVATE|MAP_DENYWRITE, 3, 0) = 0x7f1287054000
 nprotect(0x7f128723b000, 2097152, PROT_NONE) = 0
nmap(0x7f128743b000, 24576, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_FIXED|MAP_DENYWRITE, 3, 0x1e7000) = 0x7f128743b000
nmap(0x7f1287441000, 15072, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_FIXED|MAP_ANONYMOUS, -1, 0) = 0x7f1287441000
arch_prctl(ARCH_SET_FS, 0x7f12876554c0) = 0
 nprotect(0x7f128743b000, 16384, PROT_READ) = 0
nprotect(0x55bbea752000, 4096, PROT_READ) = 0
 protect(0x7f128766e000, 4096, PROT_READ) = 0
 unmap(0x7f1287656000, 96204)
brk(NULL)
                                                                 = 0x55bbeb485000
brk(0x55bbeb4a6000)
                                                                 = 0x55bbeb4a6000
exit group(0)
 student@pc:~/Desktop$
```

3. The output you received from running strace on your program was probably very messy. There's no way to tell which system call was used during the execution of malloc. Suggest a simple addition to your C code, such that you will be able to spot the system call used during the execution of malloc anyway. You're not allowed to add flags to strace. Your change must be made in the C code.

ננצל את העובדה ש-strace מנסה להדפיס את פקודות המערכת שהיא מקבלת ברצף ונכניס שתי פקודות strace ננצל את העובדה ש-malloc ובכך נוכל למצוא בקלות את קריאות המערכת בהם השתמשה התוכנית בעת ביצועו.

```
#include <stdio.h>
#include <stdlib.h>

int main(int argc, char* argv[]) {
    if (argc != 2)
    {
        printf("Usage: %s <number> :(\n", argv[0]);
    }
    int x = atoi(argv[1]);
    write(1, "malloc is next", 14);
    malloc(x);
    write(1, "malloc was here", 15);

return 0;
}
```

Section 2:

In the wet part of this homework, you wrote/will write a malloc() alternative that uses both sbrk() and mmap(). Your job in this section is to determine which memory functions the malloc() function that is included in your stdlib uses.

Hint: Use the program and the tools from the last section to help you out!

Which two system calls does the stdlib standard malloc() use in its implementation?
 Attach screenshots that prove your answer.

1) <u>brk</u> 2) <u>mmap</u>

Find the threshold that malloc uses to transition from using one function to the other. In other words, what is the number of bytes, after which calling malloc with that number, would result in using one system call instead of the other? Attach screenshots that prove your answer.

תשובה: 134537

התשובה הנתונה פה מתייחסת רק למקרה בו ה-malloc שעשינו הוא הראשון בתוכנית וכמות הבייטים היא הראשונה בה נשתמש ב-mmap בנוסף ל-brk. במקרה זה, מכיוון שמערכת ההפעלה דואגת ליישור הכתובת, היא תקצה יותר מקום מהדרוש – כמה שביקשנו ועוד כמות הזיכרון שמערכת ההפעלה מקצה לשימושה בכל בלוק – תקצה יותר מקום מהדרוש – כמה שביקשנו ועוד כמות הזיכרון שמערכת ההפעלה מימשנו בתרגיל. באופן כללי, ה-threshold דינמי ומשתנה במהלך ריצת התוכנית, אך מערכת ההפעלה לא תקצה פחות מ-135168 בייטים בעת שימוש ב-mmap. כל זאת למרות שב-mmap man כתוב כי ה-threshold מוגדר להיות 128KB=131072 דיפולטיבית.