מספר זהות:

סמסטר א, מועד ב.
תאריך: 11/3/2016
שעה: 0900
משך הבחינה: 3 שעות.
חומר עזר: אסור

בחינה בקורס: **מבוא למערכות הפעלה**

מרצה: ד"ר כרמי מרימוביץ
מתרגל: מר צבי מלמד

מדבקית
ברקוד

הנחיות:

טופס הבחינה כולל 17 עמודים (כולל עמוד זה).
תשובות צריכות לכלול הסבר.
כתיבת תשובות עמומות תוריד נקודות.
כתיבת תשובות (או חלקן) שלא קשורות לשאלות תוריד נקודות.
יש לענות בשטח המוקצה לכך.

בהצלחה!

1. (30 נק') בשאלה זו סביבת העבודה הינה קרנל לא ידוע שרץ על מעבד 64־ביטים שהינו הכללה פשוטה של פנטיום־32. ה־mmu עובד וניתן להתעלם מיחידת הסגמנטציה. כתובות וירטואליות ופיזיות הן ברוחב 64 ביטים.

במעבד זה ה־mmu עובד באופן הבא. גודל דף (פיזי ולוגי) הינו 8KB. כאשר מתרגמים כתובת וירטואלית לפיזית ישנן שלוש רמות תרגום (ולא שתיים כמו בפנטיום). לכן, בכתובת וירטואלית יש שלושה שדות אינדקסים, כל אחד ברוחב עשרה ביטים, והיסט ברוחב 13 ביטים. לא נעשה שימוש ב־21 הביטים העליונים של כתובת וירטואלית.

כל כניסה בטבלת התירגום, בכל רמה שהיא, היא בת שמונה בתים. 13 הביטים הימניים הם דגלים, כאשר הביט הימני ביותר הוא ה־valid. בשאר הדגלים אפשר להניח שערך 0 הוא תקין. 51 הביטים השמאליים הם מספר דף פיזי. (שימו לב שלמעט מספר הביטים השונה, המבנה זהה לפנטיום!)

האוגר cr3 מכיל את הכתובת הפיזית (64 ביטים!) של טבלת התירגום החיצונית, שכרגיל, נמצאת על גבול דף.

בקרנל המדובר, המשתנה הגלובלי pgdir מכיל את הכתובת הוירטואלית של הטבלה החיצונית שבתוקף, כלומר שהכתובת הפיזית שלה ב־cr3.

עליכם לכתוב פונקציה שחתימתה היא `int mmu(int *tbl, int va, int *pa)`

פונקציה זו תתרגם את הכתובת הוירטואלית va לכתובת פיזית בעזרת טבלת התירגום שנתונה על־ידי tbl. במידה והתרגום נכשל על הפונקציה להחזיר 0. אחרת עליה להחזיר 1. במידה והתירגום הצליח יש להציב ל־pa* את הכתובת הפיזית שחושבה.

לא ניתן להניח שהטבלה tbl בתוקף.

ניתן להניח שבטבלה החיצונית שבתוקף יש מספר כניסות פנויות.

int ב־C במחשב זה הוא ברוחב 64 ביטים.

אלגנטיות תילקח בחשבון בניקוד.

2. (20 נק') סביבת שאלה זו היא xv6 ב־kernel-mode. עליכם לספק שירותי mutual-exclusion (להלן mutex) ל־user-mode. ממשו את קריאות המערכת:

```
int sys_acquire(int id);
int sys_release(void);
```

id הוא בטווח 0–99 ומזהה את מספר ה־mutex אותו מנסה תהליך לרכוש. תהליך יכול לרכוש mutex אחד לכל היותר ברגע נתון. (לכן ל־sys_release אין פרמטרים.)

שימו לב ש־spinlock מספק שירותי mutex לקרנל.

אלגנטיות תילקח בחשבון בניקוד.

3. (35 נק') סביבת שאלה זו היא linux ב־user-mode.

נתונות בתכנית ההגדרות הבאות:

```
#define MAX_LEAF_PROCESSES 100
#define SILVER_NUM 13
```

נתונות וממומשות, הפונקציות הבאות:

```
double get_random();          // מחזירה מספר אקראי מטיפוס דאבל
int is_silver(double num);    // returns 1 if the num is the silver
                              // value or 0 otherwise
int sleep_random();           // התהליך הולך "לישון" למשך זמן אקראי
char set_my_class();          // returns either 'A' or 'B' based
                              // on the value on some criteria.  We don't
                              // know the internal algorithm.
```

כיתבו תכנית המבצעת את ההתנהגות הבאה:

בהתחלה תהליך ההורה (רמה 0) יוצר שני צאצאים (רמה 1).

כל אחד מהם יוצר MAX_LEAF_PROCESSES של תהליכים (רמה 2, נכדים).

מיד לאחר מכן, תהליך צאצא (רמה 1) מסתיים מבלי להמתין לאירוע כלשהו.

התהליכים נכדים (רמה 2), מבצעים את הפעילות הבאה.

תחילה כ"א מהם מברר לאיזה CLASS הוא שייך, על ידי קריאה לפונקציה set_my_class. לאחר מכן, כל צאצא מגריל מספר ע"י קריאה לפונקציה get_random() ובודק אם המספר הזה הוא "מספר הכסף" ע"י קריאה לפונקציה is_silver.

אם הערך המוחזר הוא 1 (true) אזי הצאצא מסתיים. אחרת, הצאצא מבצע קריאה לפונקציה sleep_random, וחוזר חלילה.

כאשר נכד־עלה מסוים הגריל מספר SILVER זהו סימן שהוא צריך להסתיים, אבל יחד אתו גם שאר הנכדים־עלים מאותו CLASS. כלומר, הנכדים־עלים מאותו CLASS - אם הם בדיוק ישנים, יסיימו את שנתם, אבל יפסיקו להגריל ויסתיימו. לפני שהוא יסתיים, כל תהליך־נכד יעביר את המידע "הדרוש" להורה. (ראה להלן)

כאשר כל התהליכים־נכדים הסתיימו ההורה ידפיס את הממוצע של ההגרלות בכל CLASS. הממוצע הוא סה"כ סכום ההגרלות שבוצעו באותה מחלקה לחלק במספר ההגרלות.

לדוגמא, אם במחלקה A היו רק שני תהליכים, P1, P2 וההגרלות היו:

```
P1:  2.0, 10.5, 12.5
P2:  11.4
```

אזי ההדפסה של ההורה בנוגע למחלקה הזאת תהיה:

```
Average number in class A is:  9.1
```

הנחיות / הבהרות:

(א) מירב הנקודות יינתנו למנגנונים פשוטים וקלים למימוש!

(ב) ידוע שבמערכת הנתונה יש עלות גבוהה מאוד לקריאות read() ו־write(). לפיכך אם אתם בוחרים בפתרון שמצריך שימוש בפונקציות אלו, עליכם להביא למינימום את מספר הקריאות לפונקציות אלו.

(ג) אין להתייחס למקרים של overflow.

(א) הסבירו בקצרה – 2־3 משפטים לכל היותר, איך עובר המידע מתהליכי הנכדים להורה הראשי.

(ב) הסבירו איך גורמים לצאצאים־עלים להסתיים.

(ג) ממשו את התכנית.

הערה: אין צורך לכתוב משפטי #include. כמו כן, לא צריך לבדוק הצלחה של פונקציות מערכת, כמו fork או open.

4. (15 נק') סביבת שאלה זו היא linux ב־user-mode. נתונה התוכנית הבאה.

```
1 #include <unistd.h>
2 #include <stdio.h>
3
4
5 int count = 0;
6
7 int main()
8 {
9   int pid=getpid();
10  while (fork() && fork () && (pid=fork()))
11  {
12    printf("HELLO\n");
13    count++;
14    break;
15  }
16  count++;
17  printf("AFTER pid=%d count=%d\n", getpid(), count);
18  if (pid==getpid()) printf("SAME\n");
19  return (getpid() % 255);
20 }
21
```

הניחו שמספר התהליך הראשון (ההורה הראשון) הוא 100, וכל תהליך שנוצר מקבל מספר עוקב (כלומר ה־fork הראשון יצור תהליך חדש שמספרו 101).

(א) כמה תהליכים נוצרים?

(ב) כמה פעמים מודפסת ההודעה HELLO (שורה 12)?

(ג) בסעיף זה הניחו ששורה 12 הוכנסה להערה. מהו הפלט של התכנית? (אם אפשרי מספר פלטים, תארו רק אחד מהם):

(ד) הריצו את התכנית מתוך ה־shell וכשהיא מסתיימת הריצו את הפקודה: `echo $?` (המשתנה `$?` מכיל את הערך המוחזר של התכנית). האם ניתן לדעת בוודאות מהו

הערך שהודפס? אם כן מהו? אם ישנן מספר אפשרויות ציינו מהן. נמקו בקצרה את התשובה.

```
#define PDX(va)         (((uint)(va) >> PDXSHIFT) & 0x3FF)
#define PTX(va)         (((uint)(va) >> PTXSHIFT) & 0x3FF)
#define PTXSHIFT        12      // offset of PTX in a linear address
#define PDXSHIFT        22      // offset of PDX in a linear address
#define PTE_P           0x001   // Present
#define PTE_W           0x002   // Writeable
#define PTE_U           0x004   // User
#define PTE_ADDR(pte)   ((uint)(pte) & ~0xFFF)

struct spinlock {
  uint locked;       // Is the lock held?

  // For debugging:
  char *name;        // Name of lock.
  struct cpu *cpu;   // The cpu holding the lock.
  uint pcs[10];      // The call stack (an array of program counters)
                     // that locked the lock.
};

void acquire(struct spinlock *lk)
{
  pushcli(); // disable interrupts to avoid deadlock.
  if(holding(lk))
    panic("acquire");

  while(xchg(&lk->locked, 1) != 0)
    ;


  lk->cpu = cpu;
  getcallerpcs(&lk, lk->pcs);
}

// Release the lock.
void release(struct spinlock *lk)
{
  if(!holding(lk))
```

```
    panic("release");

  lk->pcs[0] = 0;
  lk->cpu = 0;
  xchg(&lk->locked, 0);

  popcli();
}

void sleep(void *chan, struct spinlock *lk)
{
  if(proc == 0)
    panic("sleep");

  if(lk == 0)
    panic("sleep without lk");

  if(lk != &ptable.lock){  //DOC: sleeplock0
    acquire(&ptable.lock);  //DOC: sleeplock1
    release(lk);
  }

  proc->chan = chan;
  proc->state = SLEEPING;
  sched();

  proc->chan = 0;

  if(lk != &ptable.lock){  //DOC: sleeplock2
    release(&ptable.lock);
    acquire(lk);
  }
}

static void wakeup1(void *chan)
{
  struct proc *p;
```

```
  for(p = ptable.proc; p < &ptable.proc[NPROC]; p++)
    if(p->state == SLEEPING && p->chan == chan)
      p->state = RUNNABLE;
}

void wakeup(void *chan)
{
  acquire(&ptable.lock);
  wakeup1(chan);
  release(&ptable.lock);
}
```

```
WAIT(2)                    Linux Programmer's Manual                    WAIT(2)

NAME
       wait, waitpid, waitid - wait for process to change state

SYNOPSIS
       #include <sys/types.h>
       #include <sys/wait.h>

       pid_t wait(int *status);

       pid_t waitpid(pid_t pid, int *status, int options);

       int waitid(idtype_t idtype, id_t id, siginfo_t *infop, int options);
```

```
DESCRIPTION
       All of these system calls are used to wait for state changes in a child
       of  the  calling  process, and obtain information about the child whose
       state has changed.  A state change is considered to be: the child  ter-
       minated; the child was stopped by a signal; or the child was resumed by
       a signal.  In the case of a terminated child, performing a wait  allows
       the  system  to  release  the resources associated with the child; if a
       wait is not performed, then the terminated child remains in a  "zombie"
       state (see NOTES below).

       If  a  child has already changed state, then these calls return immedi-
       ately.  Otherwise they block until either a child changes  state  or  a
       signal  handler interrupts the call (assuming that system calls are not
       automatically restarted using the SA_RESTART flag of sigaction(2)).  In
       the  remainder  of this page, a child whose state has changed and which
       has not yet been waited upon by one of these  system  calls  is  termed
       waitable.
```

```
     WEXITSTATUS(status)
            returns  the  exit  status  of  the child.  This consists of the
            least significant 8 bits of the status argument that  the  child
            specified  in  a  call to exit(3) or _exit(2) or as the argument
            for a return statement in main().  This  macro  should  only  be
            employed if WIFEXITED returned true.
```

```
FLOCK(2)                    Linux Programmer's Manual                   FLOCK(2)

NAME
       flock - apply or remove an advisory lock on an open file

SYNOPSIS
       #include <sys/file.h>

       int flock(int fd, int operation);

DESCRIPTION
       Apply or remove an advisory lock on the open file specified by fd.  The
       argument operation is one of the following:

           LOCK_SH  Place a shared lock.  More than one  process  may  hold  a
                    shared lock for a given file at a given time.

           LOCK_EX  Place  an  exclusive  lock.   Only one process may hold an
                    exclusive lock for a given file at a given time.

           LOCK_UN  Remove an existing lock held by this process.

       A call to flock() may block if an incompatible lock is held by  another
       process.   To  make  a  nonblocking request, include LOCK_NB (by ORing)
       with any of the above operations.

       A single file may not simultaneously have  both  shared  and  exclusive
       locks.

       Locks  created by flock() are associated with an open file table entry.
       This means that duplicate file descriptors (created  by,  for  example,
       fork(2)  or  dup(2)) refer to the same lock, and this lock may be modi-
       fied or released using any of these descriptors.  Furthermore, the lock
       is  released  either  by  an explicit LOCK_UN operation on any of these
       duplicate descriptors, or when all such descriptors have been closed.

       If a process uses open(2) (or similar) to obtain more than one descrip-
       tor  for  the same file, these descriptors are treated independently by
       flock().  An attempt to lock the file using one of these file  descrip-
       tors  may  be  denied  by  a  lock that the calling process has already
       placed via another descriptor.
```

```
PIPE(2)                    Linux Programmer's Manual                    PIPE(2)

NAME
       pipe, pipe2 - create pipe

SYNOPSIS
       #include <unistd.h>

       int pipe(int pipefd[2]);

       #define _GNU_SOURCE             /* See feature_test_macros(7) */
       #include <unistd.h>

       int pipe2(int pipefd[2], int flags);

DESCRIPTION
       pipe()  creates  a pipe, a unidirectional data channel that can be used
       for interprocess communication.  The array pipefd is used to return two
       file  descriptors  referring to the ends of the pipe.  pipefd[0] refers
       to the read end of the pipe.  pipefd[1] refers to the write end of  the
       pipe.   Data  written  to  the write end of the pipe is buffered by the
       kernel until it is read from the read end of  the  pipe.   For  further
       details, see pipe(7).

       If  flags is 0, then pipe2() is the same as pipe().  The following val-
       ues can be bitwise ORed in flags to obtain different behavior:

       O_NONBLOCK  Set the O_NONBLOCK file status flag on  the  two  new  open
                   file  descriptions.   Using  this flag saves extra calls to
                   fcntl(2) to achieve the same result.

       O_CLOEXEC   Set the close-on-exec (FD_CLOEXEC) flag on the two new file
                   descriptors.   See  the  description  of  the  same flag in
                   open(2) for reasons why this may be useful.

RETURN VALUE
       On success, zero is returned.  On error, -1 is returned, and  errno  is
       set appropriately.
```

```
SEM_POST(3)                Linux Programmer's Manual                SEM_POST(3)

NAME
       sem_post - unlock a semaphore

SYNOPSIS
       #include <semaphore.h>

       int sem_post(sem_t *sem);

       Link with -lrt or -pthread.

DESCRIPTION
       sem_post()  increments  (unlocks)  the semaphore pointed to by sem.  If
       the semaphore's value consequently  becomes  greater  than  zero,  then
       another  process  or thread blocked in a sem_wait(3) call will be woken
       up and proceed to lock the semaphore.
```

```
NAME
       sem_getvalue - get the value of a semaphore

SYNOPSIS
       #include <semaphore.h>

       int sem_getvalue(sem_t *sem, int *sval);

       Link with -lrt or -pthread.

DESCRIPTION
       sem_getvalue() places the current value of the semaphore pointed to sem
       into the integer pointed to by sval.
```

```
NAME
       sem_wait, sem_timedwait, sem_trywait - lock a semaphore

SYNOPSIS
       #include <semaphore.h>

       int sem_wait(sem_t *sem);

       int sem_trywait(sem_t *sem);

       int sem_timedwait(sem_t *sem, const struct timespec *abs_timeout);

       Link with -lrt or -pthread.

   Feature Test Macro Requirements for glibc (see feature_test_macros(7)):

       sem_timedwait(): _POSIX_C_SOURCE >= 200112L || _XOPEN_SOURCE >= 600

DESCRIPTION
       sem_wait()  decrements (locks) the semaphore pointed to by sem.  If the
       semaphore's value is greater than zero, then  the  decrement  proceeds,
       and  the function returns, immediately.  If the semaphore currently has
       the value zero, then the call blocks until either it  becomes  possible
       to  perform the decrement (i.e., the semaphore value rises above zero),
       or a signal handler interrupts the call.

       sem_trywait() is the same as sem_wait(), except that if  the  decrement
       cannot  be immediately performed, then call returns an error (errno set
       to EAGAIN) instead of blocking.
```