ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ



ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΙΣΤΩΝ

ΕΡΓΑΣΤΗΡΙΟ ΛΕΙΤΟΥΡΓΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ – 7º ΕΞΑΜΗΝΟ

Δημητρίου Αγγελική - ΑΜ: 03117106

Τζομάκα Αφροδίτη - ΑΜ: 03117107

Ομάδα 9

2η Εργαστηριακή Άσκηση

ΕΙΣΑΓΩΓΗ - ΣΤΟΧΟΣ ΤΗΣ ΑΣΚΗΣΗΣ:

Στην εργαστηριακή άσκηση αυτή θα υλοποιηθεί οδηγός συσκευής για ένα ασύρματο δίκτυο αισθητήρων κάτω από το λειτουργικό σύστημα Linux. Το συγκεκριμένο δίκτυο διαθέτει έναν αριθμό από αισθητήρες και ένα σταθμό βάσης ο οποίος συνδέεται μέσω USB με υπολογιστικό σύστημα Linux στο οποίο και θα εκτελείται ο ζητούμενος οδηγός συσκευής - Lunix:TNG. Οι αισθητήρες αναφέρουν περιοδικά το αποτέλεσμα τριών διαφορετικών μετρήσεων: της τάσης της μπαταρίας που τους τροφοδοτεί, της θερμοκρασίας και της φωτεινότητας του χώρου όπου βρίσκονται. Ο σταθμός βάσης λαμβάνει πακέτα με δεδομένα μετρήσεων, τα οποία προωθεί μέσω διασύνδεσης USB στο υπολογιστικό σύστημα. Η διασύνδεση υλοποιείται με κύκλωμα Serial over USB, για το οποίο ήδη ο πυρήνας διαθέτει ενσωματωμένους οδηγούς, οπότε τα δεδομένα όλων των μετρήσεων όλων των αισθητήρων εμφανίζονται σε μία εικονική σειριακή θύρα, /dev/ttyUSB1.

Για την προσομοίωση της παραπάνω λειτουργίας για τον καθένα μας, σε μηχάνημα της σχολής εγκαταστάθηκε εξυπηρετητής TCP/IP, ο οποίος εκπέμπει συνεχώς μετρήσεις. Ο εξυπηρετητής αυτός είναι προσβάσιμος μέσω της θύρας TCP 49152. Το utopia.sh, ανακατευθύνει αυτόματα την πρώτη σειριακή θύρα του εικονικού μηχανήματος QEMU-KVM (/dev/ttyS0) στη θύρα TCP 49152. Οπότε στο δικό μας εικονικό μηχάνημα οι μετρήσεις εμφανίζονταν στην σειριακή θύρα/dev/ttyS0, χωρίς να υπάρχει καμία πρακτική διαφορά.

Ουσιαστικά από εμάς ζητείται η κατασκευή του οδηγού συσκευής χαρακτήρων (character device driver), ο οποίος θα λαμβάνει τα δεδομένα των μετρήσεων από το δίκτυο αισθητήρων και θα τα εξάγει στο χώρο χρήστη σε διαφορετικές συσκευές, ανάλογα με το είδος της μέτρησης και τον αισθητήρα απ' όπου προέρχεται. Το αρχείο lunix_dev_nodes.sh υλοποιεί τη δημιουργία των 16(αισθητήρες)*3(μετρήσεις) = 48 συσκευών που χρειαζόμαστε (π.χ., για τον πρώτο αισθητήρα οι /dev/lunix0-batt, /dev/lunix0-temp και /dev/lunix0-light).

ΠΕΡΙΓΡΑΦΗ ΤΩΝ ΑΝΩΤΕΡΩΝ ΕΠΙΠΕΔΩΝ

Ξεκινάμε εισάγοντας το module του LunixTNG στον πυρήνα μέσω της κλήσης της insmod και εκτελώντας την εντολή dmesg παρατηρούμε ότι καλείται η συνάρτηση lunix_module_init() του αρχείου lunix-module.c. Η ίδια αρχικοποιεί ορισμένες δομές δεδομένων όπως η lunix_sensor_struct, για την οποία δεσμεύει ένα πίνακα *lunix_sensors 16 τέτοιων δομών (όσοι και οι αισθητήρες) για γράψιμο και διάβασμα των κατάλληλων δεδομένων από όλα τα επιμέρους αρχεία. Στη συνέχεια καλούνται οι lunix_protocol_init(), lunix_sensor_init(), lunix_ldisc_init(), lunix_chrdev_init() οι οποίες εκτελούν αρχικοποιήσεις σχετικές με τις δομές που περιλαμβάνουν.

Επόμενο βήμα είναι η δημιουργία των συσκευών μέσω του lunix_dev_nodes.sh και προσάρτηση της θύρας ttyS0 μέσω του lunix-attach. Ξανατρέχουμε dmesg και αυτήν την φορά βλέπουμε την κλήση της συνάρτησης lunix_ldisc_receive() που βρίσκεται αντίστοιχα στο αρχείο lunix-ldisc.c. Н ίδια στο σώμα της καλεί lunix protocol received buff και εκείνη σειρά με τη της την lunix_protocol_update_sensors με σκοπό την πακετοποίηση των raw δεδομένων σύμφωνα με τις συμβάσεις του πρωτοκόλλου και την ανανέωση με αυτά των δομών των αισθητήρων. Για το λόγο αυτό καλείται η lunix_sensor_update και τέλος η lunix_chrdev_update διαβάζει τα σωστά δεδομένα από το πεδίο value των μετρήσεων των αισθητήρων.

ΥΛΟΠΟΙΗΣΗ ΣΥΣΚΕΥΗΣ ΧΑΡΑΚΤΗΡΩΝ LUNIX - TNG

Παρακάτω παρουσιάζονται οι συναρτήσεις του αρχείου lunix-chrdev.c οι οποίες απαιτούσαν την δική μας προσθήκη κώδικα για την ορθή λειτουργία της συσκευής. Σκοπός ουσιαστικά είναι η σωστή εξαγωγή των δεδομένων από τους buffers των αισθητήρων προς την εφαρμογή του χρήστη φροντίζοντας παράλληλα τόσο για την ταυτόχρονη πρόσβαση από πολλές διεργασίες όσο και για τον περιορισμό των δικαιωμάτων ανά αρχείο, ανά αισθητήρα, ανά μέτρηση.

Σημειώνεται πώς τα σχόλια που ξεκινάνε με «//» είναι δική μας προσθήκη και ξεχωρίζουν από τα δοθέντα που ξεκινάνε με «/*». Επεξηγούν τη λειτουργία και το σκοπό κάθε επιμέρους εντολής που χρησιμοποιήθηκε.

ο lunix_chrdev_init(): Η συνάρτηση αυτή χρησιμοποιήθηκε για την εγγραφή και την προσθήκη κάθε συσκευής στον οδηγό μέσω των συναρτήσεων register_chrdev_region και cdev_add αντίστοιχα. Η πρώτη καταχωρεί τη συσκευή δεσμεύοντας έναν major number, μια σειρά minor numbers και ένα όνομα για αυτήν. Η δεύτερη προσθέτει στο σύστημα και θέτει σε λειτουργία τη συσκευή. Προφανώς υπάρχει και η αντίστοιχη συνάρτηση destroy που περιέχει τα cdev_delete και uregister_chrdev_region απλά δεν χρειάστηκε κάποια προσθήκη από εμάς.

```
lunix_chrdev_init(void)
dev no = MKDEV(LUNIX CHRDEV MAJOR, 0);
```

ο lunix_chrdev_open(): Η open είναι η πρώτη ρουτίνα που υλοποιήσαμε και στην οποία δείχνει το struct file_operations. Γενικά, ελέγχει για τυχόν σφάλματα που έχουν να κάνουν με το εκάστοτε αρχείο της συσκευής, αρχικοποιεί όλα τα απαραίτητα πεδία και δεσμεύει την κατάλληλη μνήμη. Σημειώνεται ότι στο πεδίο private_data του ανοιχτού αποθηκεύεται όλη η πληροφορία για την αρχική κατάσταση της συσκευής χαρακτήρων η οποία μετέπειτα θα περάσει και στα διάφορα system calls, κάτι το οποίο περιγράφεται εκτενώς στα σχόλια.

```
static int lunix chrdev open(struct inode *inode, struct file *filp)
{
    /* Declarations */
    /* ? */
    int ret;
    int minor num;
    struct lunix chrdev state struct *prst;
```

```
((ret = nonseekable open(inode, filp)) < 0) //Nonseekable open clears the bits that
```

ο lunix_chrdev_release(): Η συνάρτηση αυτή απελευθερώνει τον δείκτη private_data κατά το κλείσιμο του αρχείου καθώς δεν θα είναι πλέον open file και δεν θα θέλαμε ξεκρέμαστο pointer.

```
static int lunix chrdev release(struct inode *inode, struct file *filp)
{
    /* ? */
    kfree(filp->private_data); //Free the private_data pointer
    return 0;
}
```

ο lunix_chrdev_ioctl(): Η προαιρετική αυτή συνάρτηση υλοποιεί λειτουργίες input/output που δεν υποστηρίζονται από τις κλασσικές κλήσεις συστήματος. Μιας και δεν υλοποιήθηκε, οφείλει να επιστρέφει invalid argument. Το ίδιο ισχύει και για την lunix_chrdev_mmap().

```
static long lunix_chrdev_ioctl(struct file *filp, unsigned int cmd, unsigned long arg)
{
   /* Why? */
   //If the device doesn't provide an ioctl method
   //the system call returns an error for any request that isn't predefined
   return -EINVAL; //invalid argument
}
```

lunix_chrdev_read(): Εδώ υλοποιείται η επιστροφή των δεδομένων στον χρήστη. Αρχικά ανακτούνται οι πληροφορίες τόσο για την παρούσα κατάσταση μέσω του πεδίου private_data του δοθέντος ανοικτού αρχείου, όσο και για τον αντίστοιχο αισθητήρα. Στη συνέχεια αφού εξασφαλιστεί ότι μόνο μία διεργασία εκ του συνόλου των διεργασιών που μοιράζονται το ίδιο struct file (πχ πατέρας - παιδί) θα αποκτήσει πρόσβαση στα δεδομένα, δηλαδή θα αποκτήσει τον σημαφόρο, ελέγχεται η θέση από την οποία θα γίνει το read. Αν βρισκόμαστε στην αρχή θα πρέπει να κάνουμε τους απαιτούμενους ελέγχους για την ύπαρξη ή όχι καινούριων δεδομένων φροντίζοντας τη συμπεριφορά των διεργασιών στην δεύτερη περίπτωση (απελευθέρωση σημαφόρου, εξασφάλιση blocking λειτουργίας, διακοπτόμενος ύπνος, ανάκτηση σημαφόρου). Σημειώνεται ότι όσο κοιμούνται οι διεργασίες με την wait_event_interruptible, μέσω της wake_up_interruptible της lunix_sensor_update() ξυπνάνε όλες όσες βρίσκονται στο waiting queue και ελέγχουν την συνθήκη ανανέωσης δεδομένων. Αν ισχύει μόνο η σωστή διεργασία θα λάβει εν τέλει τον σημαφόρο μέσω της down_interruptible. Αν βρισκόμαστε στο τέλος τότε επιστρέφουμε το 0 ενώ διαφορετικά, ξεκινάμε τη διαδικασία του διαβάσματος. Πιο συγκεκριμένα, καθορίζεται ο αριθμός δεδομένων που θα διαβαστούν, τα δεδομένα αυτά μεταφέρονται στους buffers του χρήστη και ανανεώνεται κατάλληλα το offset μέσα στον buffer δεδομένων της κατάστασης που διαβάζουμε. Τιμή επιστροφής τίθεται ο αριθμός των δεδομένων που διαβάστηκαν και αν φτάσαμε στο τέλος επιστρέφουμε το offset στην αρχή.

```
if (*f_pos == 0)
    while (lunix chrdev state update(state) == -EAGAIN) //EAGAIN = no available data, try
        up(&state->lock);
        if (filp->f_flags & O_NONBLOCK) //if in non-blocking mode, try again
    return -EAGAIN;
        if (wait event interruptible(sensor->wq, lunix chrdev state needs refresh(state)))
if (*f pos >= state->buf lim)
   cnt = state->buf lim - *f pos;
up(&state->lock);
```

ο lunix_chrdev_state_needs_refresh(): Πρόκειται για βοηθητική συνάρτηση ώστε να γνωρίζουμε εάν τα δεδομένα μέτρησης της ιδιωτικής κατάστασης που αποθηκεύτηκε στο πεδίο private_data του ανοιχτού αρχείου πρέπει να ανανεωθούν ή όχι. Επιστρέφει 1 εάν ναι και μηδέν αν όχι. Ελέγχει αν έχει συμβεί ανανέωση των δεδομένων στη δομή του αισθητήρα της κατάστασης σε

μεταγενέστερη χρονική στιγμή από αυτήν που είναι αποθηκευμένη στο πεδίο buf_timestamp. Σημειώνεται πως το πεδίο last_update αλλάζει λόγω καινούριας μέτρησης των αισθητήρων, ενώ το buf_timestamp ανανεώνεται από εμάς. Δεδομένου ότι η συνάρτηση αυτή διαβάζει απευθείας δεδομένα του αισθητήρα, θα πρέπει να καλείται μόνο όταν η διεργασία που την καλεί έχει πάρει το spinlock του αισθητήρα.

```
static int lunix_chrdev_state_needs_refresh(struct lunix_chrdev_state_struct *state)
{
    struct lunix sensor struct *sensor;

    WARN_ON(!(sensor = state->sensor));
    //If the condition above is true (error) WARN() prints the contents of the registers and the stack Trace.

    //Function returns 1 if state needs refreshing or 0 if not
    //The state needs to be refreshed only when an update to the measurement
    //data has happened later than the timestamp of the given state
    return sensor->msr_data[state->type]->last_update > state->buf_timestamp;
}
```

ο lunix_chrdev_state_update(): Πρόκειται για τη συνάρτηση που ανανεώνει τα πεδία της κατάστασης του οδηγού συσκευής χαρακτήρων όταν αυτό είναι αναγκαίο (νέα διαθέσιμα δεδομένα). Λόγω της λειτουργίας της η update οφείλει να καλείται από διεργασία αναγνώστη μόνο αφού αποκτήσει σημαφόρο. Η ανανέωση των δεδομένων χωρίζεται σε δύο στάδια, την απόκτηση τους από τον αισθητήρα (επεξεργασμένα δεδομένα από την update του lunix protocol) και την αποθήκευση τους στον buffer της ιδιωτικής κατάστασης. Δεδομένου ότι η needs_refresh επιστρέψει 1, η ανάκτηση δεδομένων πρέπει να γίνει γρήγορα κρατώντας το spinlock του αισθητήρα ώστε να μην ενοχληθεί από κάποια διακοπή. Όταν τελειώσει αυτή η διαδικασία γυρνάμε στο ίδιο interrupt state που είχαμε πριν αποκτηθεί το spinlock και πλέον μπορούμε χωρίς βιασύνη να κάνουμε format στα δεδομένα με τον τρόπο που αναλυτικά περιγράφεται στα σχόλια του κώδικα.

ΕΛΕΓΧΟΣ ΟΡΘΗΣ ΛΕΙΤΟΥΡΓΙΑΣ

Αρχικά ελέγχθηκε η ορθή λειτουργία του οδηγού με χρήση της εντολής cat από διαφορετικά τερματικά, διαβάζοντας ταυτόχρονα και από ίδιους αισθητήρες και βλέπουμε ότι τα αποτελέσματα είναι τα αναμενόμενα. Δηλαδή κάθε διεργασία έχει πρόσβαση στις μετρήσεις και δεν παρουσιάζονται προβλήματα/παγώματα στην εκτέλεση του προγράμματος.

Για τον περαιτέρω έλεγχο της ορθότητας του οδηγού, φτιάξαμε και ένα userspace program το οποίο ουσιαστικά κάνει fork ένα παιδί και διαβάζει τα δεδομένα των μετρήσεων της συσκευής που δίνεται στο input. Παρατηρούμε πως αν και ο πατέρας και το παιδί έχουν πρόσβαση στο ίδιο fd και προσπαθούν να διαβάσουν και οι δύο από αυτό τα αποτελέσματα και πάλι είναι σωστά. Αναλυτικότερα, καταφέρνουν να διαβάσουν και οι δύο μετρήσεις εναλλάξ χωρίς να μπερδεύονται.

Τα αποτελέσματα αυτά φαίνονται στις εικόνες παρακάτω μαζί με τον κώδικα του userspace προγράμματος.

```
Debtan CMU/Linux comes with ABSOLUTELY NO MARRANTY, to the extent permitted by applicable law.

You have new mall.
Last login: Tue Dec 8 17:3812 2020 from 10.0.2.2

user@utopla:-
4.190
4.190
4.191
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4.190
4
```

Αποτελέσματα cat

```
post of the process of the process
```

Αποτελέσματα μέσω του προγράμματος χρήστη

```
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <sys/types.h>
#include <sys/stat.h>
#include <sys/wait.h>
#include <sys/ioctl.h>
#include <fcntl.h>
#include <string.h>
#include "lunix-chrdev.h"
int main(int argc, char **argv)
    int fd, status;
   pid t pid;
   size_t bytes_read2;
   char buffer2[7];
    if (argc != 2)
        fprintf(stderr, "Usage: ./test /dev/lunixX-XXXX\n");
    fd = open(argv[1], O_RDONLY);
    if (fd < 0)
        perror(argv[1]);
    printf("Father pid: %d\n", getpid());
    pid = fork();
    if (pid == 0)
        printf("Child pid: %d\n", getpid());
    do
        bytes_read2 = read(fd, &buffer2, 7);
        if (bytes_read2 < 0)</pre>
            perror("read");
            exit(1);
        printf("%s , %d\n", buffer2, getpid());
    } while (bytes_read2 > 0);
    //Unreachable
    perror("reached non reachable point, internal error");
    exit(0);
    wait(&status);
    return 0;
```