Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Εργαστήριο Λειτουργικών Συστημάτων

2η Άσκηση

«Σχεδιασμος και υλοποιηση υποδομης σημειωσης διεργασιων στον πυρηνα του Linux»

Τελική Αναφορά

25 Ιουλίου 2012

Διονύσης Ζήνδρος <dionyziz@gmail.com> (03106601)
Δημήτρης Κονόμης <dim.konomis@gmail.com> (03107015)

Σύνοψη

Το σύστημα «κατάρων» είναι ένα σύστημα σημείωσης διεργασιών στον πυρίνα του Linux το οποίο επιτρέπει την εφαρμογή ονομασμένων σημειώσεων σε διεργασίες. Οι κατάρες είναι ονομασμένες και μπορούν να αλλάξουν με προσθήκη νέων με την απλή συγγραφή και σύνδεση του αντίστοιχου module στον πυρίνα, ενώ ενεργοποιούνται και απενεργοποιούνται στο χρόνο εκτέλεσης από το userland. Επιτρέπεται επίσης η καθολική ενεργοποίηση και απενεργοποίησή τους. Στην παρούσα αναφορά παρουσιάζουμε τις τεχνικές λεπτομέρειες υλοποίησης του συστήματος.

Η κατάρα no-fs-cache είναι μία κατάρα που υλοποιήθηκε αξιοποιώντας την υποδομή κατάρων που αναπτύχθηκε. Η συγκεκριμένη κατάρα δεν επιτρέπει στη διεργασία στην οποία εφαρμόζεται να αξιοποιεί το σύστημα caching αρχείων στη μνήμη του πυρίνα, με αποτέλεσμα η επαναλαμβανόμενη πρόσβαση σε αρχεία να χρησιμοποιεί κάθε φορά απευθείας το δίσκο.

Ο κώδικας της υλοποίησης περιλαμβάνεται στο παράρτημα.

Διεπαφή με το χώρο χρήστη

Η επικοινωνία ανάμεσα στο userland και στον πυρίνα γίνεται μέσω της κλήσης συστήματος curse():

```
long curse(long call, curse_id_t curse_id, pid_t pid, void* addr);
```

Η συγκεκριμένη κλήση αναλαμβάνει την επικοινωνία ανάμεσα στο userland και τον πυρίνα για όλες τις ανάγκες που αφορούν την σημείωση διεργασιών. Η παράμετρος call καθορίζει την ενέργεια που πρέπει να γίνει και οι υπόλοιπες παράμετροι περνούν τα απαραίτητα δεδομένα. Η call μπορεί να πάρει διαφορετικές τιμές και αναλόγως εκτελείται η αντίστοιχη λειτουργία. Οι παράμετροι `curse_id`, `pid` και `addr` χρησιμοποιούνται για να περάσουν δεδομένα από το χώρο χρήστη στο χώρο πυρίνα ή αντίστροφα.

Σε κάθε κλήση, ενδέχεται ορισμένες παράμετροι να μην χρησιμοποιούνται. Για παράδειγμα, στην κλήση `CURSE_CMD_GET_CURSES_LIST` (βλ. παρακάτω) δεν χρησιμοποιούνται οι παράμετροι `pid` και `curse_id`.

Παρατίθενται οι επιτρεπόμενες κλήσεις:

`CURSE_CMD_GET_CURSES_LIST:`

Επιστρέφει τη λίστα με τις διαθέσιμες κατάρες που υποστηρίζει ο πυρίνας. Αυτές γράφονται στη διεύθυνση που περνάει ως παράμετρος μέσω της `addr`. Τα ονόματα από τις διαθέσιμες κατάρες γράφονται στο `*addr` η μία μετά την άλλη ως strings χωρισμένες με null χαρακτήρες. Στο τέλος της λίστας ακολουθεί ένας επιπλέον null χαρακτήρας που επισημαίνει το τέλος της λίστας.

`CURSE_CMD_CURSE_GLOBAL_STATUS:`

Ελέγχει αν η δεδομένη κατάρα που περνάει ως παράμετρος μέσω του `curse_id` είναι ενεργοποιημένη ή απενεργοποιημένη καθολικά.

`CURSE_CMD_CURSE_GLOBAL_ENABLE:`

Ενεργοποιεί την κατάρα που περνά ως παράμετρος `curse_id` καθολικά.

`CURSE_CMD_CURSE_GLOBAL_DISABLE:`

Απενεργοποιεί την κατάρα που περνά ως παράμετρος `curse_id` καθολικά.

`CURSE_CMD_CURSE_STATUS:`

Ελέγχει αν η δεδομένη κατάρα `curse_id` είναι ενεργοποιημένη για τη διεργασία `pid`.

`CURSE_CMD_CURSE_CAST:`

Ενεργοποιεί την κατάρα curse_id για τη διεργασία `pid`.

`CURSE_CMD_CURSE_LIFT:`

Απενεργοποιεί την κατάρα curse_id για τη διεργασία `pid`.

Ως παράδειγμα χρήσης της κλήσης συστήματος, ο χρήστης του API συμβουλεύεται να κοιτάξει τον κώδικα του libcurse.c που παρατίθεται στο παράρτημα.

Υλοποίηση

Ο καλύτερος τρόπος να περιγραφεί η υλοποίηση είναι με παράθεση του κώδικα, γι' αυτό η παρούσα περιγραφή περιορίζεται σε υψηλού επιπέδου σχόλια και ο κώδικας παρατίθεται στο παράρτημα.

Το ποιες κατάρες είναι ενεργοποιημένες καθολικά αποθηκεύεται στην καθολική μεταβλητή curses_status που είναι ένα bitmask με ένα bit ανά κατάρα. Καθώς όλες οι κατάρες είναι αρχικά ενεργοποιημένες εκ προεπιλογής, η μεταβλητή αρχικοποιείται σε άσσους:

```
static int curses_status = 0xffffffff;
```

Το ποιες κατάρες είναι απενεργοποιημένες ανά διεργασία αποθηκεύεται μέσα στο `task_struct`. Αυτό γίνεται έτσι ώστε να κληρονομούνται τα αντίστοιχα δεδομένα κατά τη διαδικασία του forking:

```
struct task_struct {
…
     unsigned int curses;
}
```

Και πάλι το συγκεκριμένο πεδίο λειτουργεί ως bitmask.

Διαχείριση δικαιωμάτων

Το σύστημα κατάρων ελέγχει τα δικαιώματα χρηστών πριν εφαρμόσει οποιαδήποτε λειτουργία. Συγκεκριμένα, η καθολική ενεργοποίηση και απενεργοποίηση κατάρων απαιτεί δικαιώματα root, ενώ για να καταραστεί κανείς μία συγκεκριμένη διεργασία απαιτείται είτε να είναι root είτε να είναι ο ιδιοκτήτης της εν λόγω διεργασίας.

Βιβλιοθήκη χώρου χρήστη

Ο ρόλος της βιβλιοθήκης είναι να προσφέρει έναν ευκολότερο μηχανισμό για την πρόσβαση στην κλήση συστήματος. Συγκεκριμένα προσφέρονται διαφορετικές συναρτήσεις για κάθε μία από τις διαφορετικές πιθανές τιμές της παραμέτρου call που περνά στο system call curse. Επιπλέον, η λίστα των κατάρων επιστρέφεται ως συνδεδεμένη λίστα και το API είναι πιο ξεκάθαρο. Δεν είναι κάτι άλλο πέρα από έναν wrapper για το system call.

Εργαλείο χώρου χρήστη

Το εργαλείο χώρου χρήστη αξιοποιεί τη βιβλιοθήκη χώρου χρήστη για να επιτρέψει λειτουργίες πάνω στις κατάρες χωρίς να χρειάζεται η συγγραφή κώδικα. Είναι ουσιαστικά ένα CLI πρόγραμμα το οποίο με απλή σύνταξη επιτρέπει στο χρήστη να παρέμβει στις κατάρες και να έχει πρόσβαση στην κλήση συστήματος με όλες τις λειτουργίες που αυτή επιτρέπει. Καλώντας `./curse --help` μπορείτε να δείτε ποιες ακριβώς κλήσεις επιτρέπονται, την σύνταξή τους και τι ενέργειες επιτελούν.

Κατάρα no-fs-cache

Η κατάρα υλοποιείται καλώντας τη συνάρτηση `fadvice` για να απελευθερωθούν τα δεδομένα που έχουν αποθηκευτεί στην cache από τη μνήμη. Αυτό γίνεται παρεμβαίνοντας στην κλήση συστήματος read και write και καλώντας τη συνάρτηση checkpoint που ελέγχει αν η κατάρα είναι ενεργοποιημένη και καλεί την `fadvice` αντίστοιχα. Για λόγους ταχύτητας, το σύστημα αποθηκεύει εσωτερικά έναν μετρητή ώστε να μην γίνεται `fadvice` πάρα πολύ συχνά, αλλά μόνο αφότου διαβαστεί ένα δεδομένο πλήθος από bytes.

Παράρτημα

Ο κώδικας παρατίθεται στο:

https://github.com/dionyziz/ntua-kernel

Στο git version history φαίνονται ακριβώς οι αλλαγές που έγιναν πάνω στον υπάρχοντα κώδικα του πυρίνα.