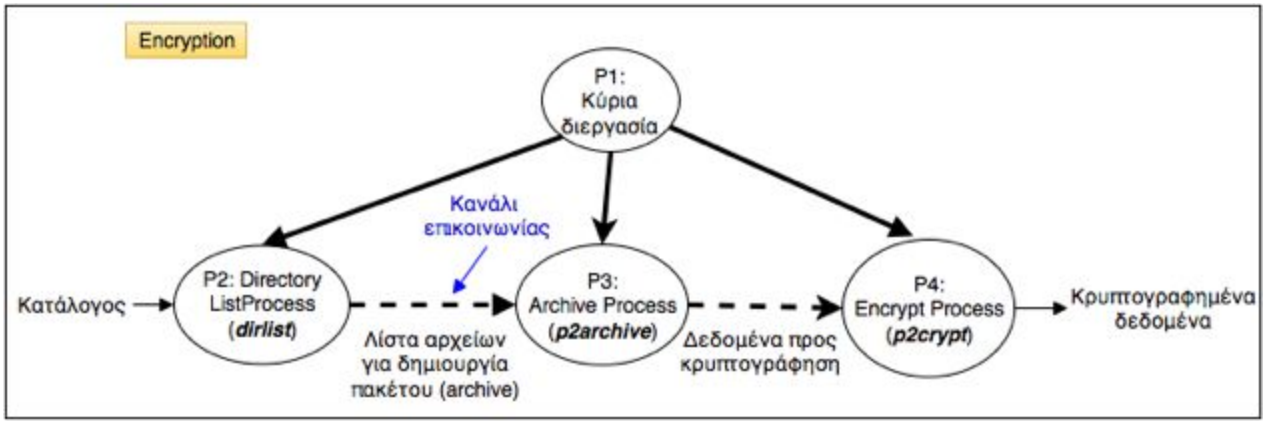
Εργασία 2 – Ροή διεργασιών για τη δημιουργία κρυπτογραφημένου archive αρχείων

Αναπτύξτε μια εφαρμογή για την ενσωμάτωση των περιεχομένων ενός καταλόγου (directory) σε ένα κρυπτογραφημένο αρχείο archive, και αντίστροφα για την εξαγωγή των αρχείων που περιέχει ένα τέτοιο κρυπτογραφημένο archive σε έναν κατάλογο. Το πρόγραμμα λαμβάνει ως ορίσματα (α) μια επιλογή -E (για archive & encrypt) ή -D (για decrypt & extract), (β) το όνομα ενός καταλόγου, (γ) ένα αλφαριθμητικό κλειδί, και (δ) το όνομα ενός αρχείου.

Διαδικασία ενσωμάτωσης και κρυπτογράφησης (επιλογή -E)

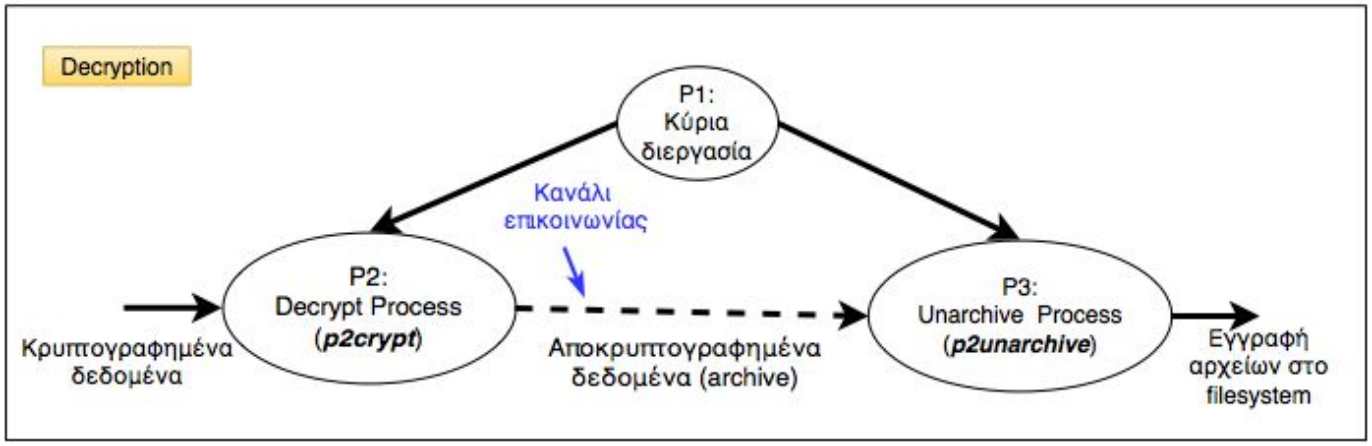
Η διαδικασία ενσωμάτωσης και κρυπτογράφησης υλοποιείται μέσα από μια ροή επεξεργασίας ως εξής:



- **Η διεργασία P1 εκτελεί το κυρίως πρόγραμμα.** Ο κατάλογος που δίνεται ως όρισμα είναι αυτός από τον οποίο η P2 θα διαβάσει τα αρχεία, ενώ στο αρχείο θα αποθηκευτούν τελικά τα κρυπτογραφημένα δεδομένα. Αν το αρχείο υπάρχει ήδη, το πρόγραμμα τερματίζει αφού πρώτα εκτυπώσει κατάλληλο μήνυμα. Αν το αρχείο δεν υπάρχει, το δημιουργεί και γράφει στην αρχή του το μοναδικό *magic number* `P2CRYPTAR` (9 bytes). Κατόπιν, φτιάχνει την **ροή επεξεργασίας**: (α) **δημιουργεί τις διεργασίες** P2, P3, P4 (περνώντας τους τα ορίσματα των προγραμμάτων που θα εκτελέσει η κάθε μια), (β) δημιουργεί τους **αγωγούς** που θα χρησιμοποιηθούν ως κανάλια επικοινωνίας, και (γ) **ανακατευθύνει τις καθιερωμένες εισόδους/εξόδους των διεργασιών**, έτσι ώστε η P2 να γράφει στον αγωγό από όπου διαβάζει η P3, η P3 να γράφει στον αγωγό από όπου διαβάζει η P4 και η P4 να γράφει στο αρχείο που κατασκεύασε η P1.
- **Η διεργασία P2 εκτελεί το πρόγραμμα dirlist** που παίρνει ως όρισμα το όνομα ενός καταλόγου και, εφόσον αυτός υπάρχει, εκτυπώνει στην καθιερωμένη έξοδο τα ονόματα των αρχείων που περιέχει ο κατάλογος (με πληροφορία path) σε ξεχωριστή γραμμή το καθένα. Κατά την εκτύπωση των περιεχομένων του καταλόγου, πρέπει να αγνοούνται όλοι οι υποκατάλογοι. [hint: man 2 stat ή man opendir]. Αν ο κατάλογος δεν υπάρχει, το πρόγραμμα τερματίζει αφού πρώτα εκτυπώσει κατάλληλο μήνυμα στην καθιερωμένη έξοδο λαθών.
- **Η διεργασία P3 εκτελεί το πρόγραμμα p2archive** που διαβάζει από την καθιερωμένη είσοδό του μια σειρά από ονόματα αρχείων, και γράφει στην καθιερωμένη έξοδο του όλη την πληροφορία που θέλουμε να εισαχθεί στο αντίστοιχο archive. Λεπτομέρειες για τη μορφή του archive δίνονται παρακάτω.
- **Η διεργασία P4 εκτελεί το πρόγραμμα p2crypt** που παίρνει ως όρισμα ένα αλφαριθμητικό κλειδί, διαβάζει από την καθιερωμένη είσοδό του δεδομένα τα οποία κρυπτογραφεί με βάση το κλειδί, και τα γράφει στην καθιερωμένη έξοδό του.

Διαδικασία αποκρυπτογράφησης και εξαγωγής (επιλογή -D)

Η διαδικασία αποκρυπτογράφησης και εξαγωγής υλοποιείται μέσα από μια ροή επεξεργασίας ως εξής:



- **Η διεργασία P1 εκτελεί το κυρίως πρόγραμμα.** Το αρχείο που δίνεται ως όρισμα περιέχει τα κρυπτογραφημένα δεδομένα, ενώ ο κατάλογος είναι αυτός στον οποίο θα εξαχθούν τα αρχεία από το archive. Αν το αρχείο δεν υπάρχει ή δεν έχει το σωστό *magic number* (`P2CRYPTAR`), το πρόγραμμα τερματίζει μετά την εκτύπωση κατάλληλου μηνύματος. Διαφορετικά, φτιάχνει την **ροή επεξεργασίας**: (α) **δημιουργεί τις διεργασίες** P2, P3 (περνώντας τους τα ορίσματα των προγραμμάτων που θα εκτελέσει η κάθε μια), (β) **δημιουργεί τους αγωγούς** που θα χρησιμοποιηθούν ως κανάλια επικοινωνίας, και (γ) **ανακατευθύνει τις καθιερωμένες εισόδους/εξόδους** των διεργασιών, έτσι ώστε η P2 να διαβάζει από το αρχείο που περιέχει τα κρυπτογραφημένα δεδομένα και να γράφει στον αγωγό από τον οποίο διαβάζει η P3.

- **Η διεργασία P2 εκτελεί το πρόγραμμα p2crypt** που παίρνει ως όρισμα το αλφαριθμητικό κλειδί κρυπτογράφησης. Το πρόγραμμα διαβάζει κρυπτογραφημένα δεδομένα από την καθιερωμένη είσοδο, τα αποκρυπτογραφεί με χρήση του κλειδιού, και τα γράφει στην καθιερωμένη έξοδο του.

- **Η διεργασία P3 εκτελεί το πρόγραμμα p2unarchive** που παίρνει ως όρισμα το όνομα ενός καταλόγου. Αν ο κατάλογος υπάρχει ήδη, το πρόγραμμα τερματίζει αφού πρώτα εκτυπώσει κατάλληλο μήνυμα στην καθιερωμένη έξοδο λαθών. Διαφορετικά, δημιουργεί τον κατάλογο. Κατόπιν, διαβάζει από την καθιερωμένη είσοδο τα περιεχόμενα ενός archive, κι εξάγει τα αρχεία που αυτό περιέχει μέσα στον κατάλογο, χρησιμοποιώντας τα ονόματα, τις χρονοσημάνσεις τελευταίας πρόσβασης και μεταβολής, και τα δικαιώματα πρόσβασης που βρίσκονται αποθηκευμένα στο archive. [Hint: man 2 utimes, man 2 chmod, man 2 stat].

Μορφή archive (μη κρυπτογραφημένη)

Ένα archive περιέχει μια σειρά από εγγραφές, μια για κάθε αρχείο που περιέχει το archive, με την εξής μορφή: **(1)** το μήκος του ονόματος του αρχείου (σε binary), **(2)** τους χαρακτήρες του ονόματος χωρίς κάποια πληροφορία μονοπατιού (σε ascii), **(3)** χρονοσήμανση τελευταίας πρόσβασης και τελευταίας μεταβολής του αρχείου (σε binary) [hint: man 2 stat], **(4)** ένα πεδίο τύπου mode_t (σε binary), όπου καταχωρείται το είδος του αρχείου και τα δικαιώματα πρόσβασης σε αυτό [hint: man 2 stat], **(5)** το μέγεθος του αρχείου (τύπος off_t σε binary), και τέλος **(6)** το περιεχόμενο του αρχείου (ως έχει, χωρίς ερμηνεία/επεξεργασία).

Κρυπτογράφηση / αποκρυπτογράφηση

Η κρυπτογράφηση των δεδομένων γίνεται συνδυάζοντας τα δεδομένα μέσω XOR με ένα αλφαριθμητικό κλειδί (byte προς byte). Η αποκρυπτογράφηση γίνεται με τον ίδιο ακριβώς τρόπο.

Όποιος έχει κέφι για κάτι παραπάνω

Επεκτείνετε την υπάρχουσα λειτουργικότητα του προγράμματος σας έτσι ώστε μετά την κρυπτογράφηση, να υπολογίζεται και να προστίθεται στο τέλος του archive ένα ψηφιακό αποτύπωμα (fingerprint), το οποίο να ελέγχεται για να επιβεβαιωθεί η ακεραιότητα του archive προτού αρχίσει η διαδικασία εξαγωγής. Ψάξτε στο διαδίκτυο για κατάλληλες fingerprinting functions και επιλέξτε κάποια που να σας αρέσει.

Γενικές παρατηρήσεις και απαιτήσεις

Αναπτύξτε κάθε ένα από τα προγράμματα (p2archive, p2crypt, p2unarchive) ως ένα αυτόνομο εκτελέσιμο, και βεβαιωθείτε για την σωστή λειτουργία του ξεχωριστά.

Το πρόγραμμα που εκτελεί η κύρια διεργασία είναι ένα, και δημιουργεί την κατάλληλη ροή επεξεργασίας ανάλογα με την επιλογή -E ή -D.

Οι διεργασίες-παιδιά **πρέπει να υφίστανται ταυτόχρονα**. Η κυρίως διεργασία δεν πρέπει να περιμένει να τερματίσει το ένα παιδί προτού ξεκινήσει το επόμενο ή να επιβάλει κάποια σειρά εκτέλεσης.

Η κύρια διεργασία πρέπει να περιμένει να τερματίσουν οι διεργασίες-παιδιά που δημιούργησε.

Μηνύματα λάθους πρέπει να εκτυπώνονται στο stderr.

Πιθανώς να σας φανεί χρήσιμο να ενσωματώσετε προσωρινά στο PATH του συστήματος σας τον κατάλογο στον οποίο βρίσκονται τα εκτελέσιμα που θα δημιουργήσετε.

Όπως πάντα, απαγορεύεται η χρήση καθολικών μεταβλητών, goto, gets. Δώστε περιγραφικά ονόματα σε συναρτήσεις και μεταβλητές, σχολιάστε τις συναρτήσεις και όσα τμήματα του κώδικα δεν είναι ξεκάθαρα, και γράψτε ευανάγνωστα με σωστή στοίχιση. Κώδικας που «δεν διαβάζεται» θα απορριφθεί, χωρίς εξέταση.

Με τα αρχεία σας πρέπει να συμπεριλάβετε κατάλληλο Makefile.

Συζήτηση/επεξήγηση εργασίας: **Πέμπτη 15/3/2018**, στην ώρα του μαθήματος Προθεσμία παράδοσης εργασίας: **Κυριακή 1/4/2018, 23:59** Οδηγίες παράδοσης εργασίας: στη σελίδα του μαθήματος.

Βασικά στάδια ανάπτυξης του κωδικα

Στάδιο 1

Γράψτε το πρόγραμμα dirlist που θα εκτελέσει η διεργασία P2 .

Για να ελέγξετε την ορθή λειτουργία του dirlist, χρησιμοποιώντας ένα κατάλογο test στον οποίο έχετε διάφορα αρχεία, εκτελέστε:

./dirlist test

Θα πρέπει να εμφανιστούν τα ονόματα αρχείων που βρίσκονται στο test, με κατάλληλη πληροφορία μονοπατιού (π.χ. test/file1), με κάθε όνομα σε ξεχωριστή γραμμή.

Στάδιο 2

Γράψτε τα προγράμματα p2archive και p2unarchive.

Για να ελέγξετε την ορθή λειτουργία της p2archive σε συνδυασμό με την dirlist εκτελέστε:

./dirlist test | ./p2archive > hw2.p2ar

Το hw2.p2ar είναι το παραγόμενο archive και μπορείτε να ελέγξετε τα περιεχόμενά του μέσω του προγράμματος xxd.

Ελέγξτε την ορθή εξαγωγή των περιεχομένων του αρχείου hw2.p2ar ως εξής:

./p2unarchive testcopy < hw2.p2ar

Μετά την εκτέλεση, ο κατάλογος testcopy θα πρέπει να περιέχει τα αρχεία που περιείχε και ο test.

Στάδιο 3

Γράψτε το πρόγραμμα p2crypt.

Ελέγξτε την ορθή κρυπτογράφηση ενός αρχείου testfile1 με κλειδί “my-secret-key” ως εξής:

./p2crypt my-secret-key < testfile1 > testfile2

Μετά την εκτέλεση, το αρχείο testfile2 περιέχει τα κρυπτογραφημένα δεδομένα, και με την σειρά του μπορεί να αποκρυπτογραφηθεί ως εξής:

./p2crypt my-secret-key < testfile2 > testfile3

Το αρχείο testfile3 πρέπει να έχει τα ίδια περιεχόμενα με το αρχείο testfile1.

Στάδιο 4

Γράψτε το πρόγραμμα που υλοποιεί την κύρια διεργασία. Το τελικό πρόγραμμα θα πρέπει να μπορείτε να το καλέσετε ως εξής:

Για την δημιουργία του κρυπτογραφημένου archive:

./hw2 -E dirname my-secret-key hw2.p2enc

όπου dirname το όνομα του καταλόγου με τα αρχεία προς εισαγωγή στο archive, my-secret-key είναι το κλειδί κρυπτογράφησης, και hw2.p2enc είναι το όνομα του τελικού κρυπτογραφημένου archive.

Για την αποκρυπτογράφηση κι εξαγωγή των περιεχομένων του κρυπτογραφημένου archive:

./hw2 -D dirname my-secret-key hw2.p2enc

όπου dirname το όνομα του καταλόγου στον οποίο θα εξαχθούν τα αρχεία, my-secret-key είναι το κλειδί κρυπτογράφησης, και hw2.p2enc το όνομα του κρυπτογραφημένου archive που πρόκειται να αποκρυπτογραφηθεί και να εξαχθούν τα περιεχόμενά του.