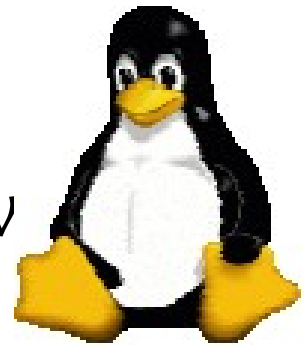




τμήμα Εφαρμοσμένης Πληροφορικής

Προγραμματισμός και Διαχείριση Συστημάτων I

Μάθημα 7ο
Εργαλεία αναζήτησης αρχείων



Κ. Μαργαρίτης, Π. Μιχαηλίδης

Γ' Εξάμηνο 2007

Εργαλεία αναζήτησης αρχείων

Σύνδεσμοι αρχείων

Απλή αναζήτηση αρχείων

Κανονικές εκφράσεις

Εκτεταμένες κανονικές εκφράσεις

Σύνθετη αναζήτηση αρχείων

Σύνδεσμοι αρχείων

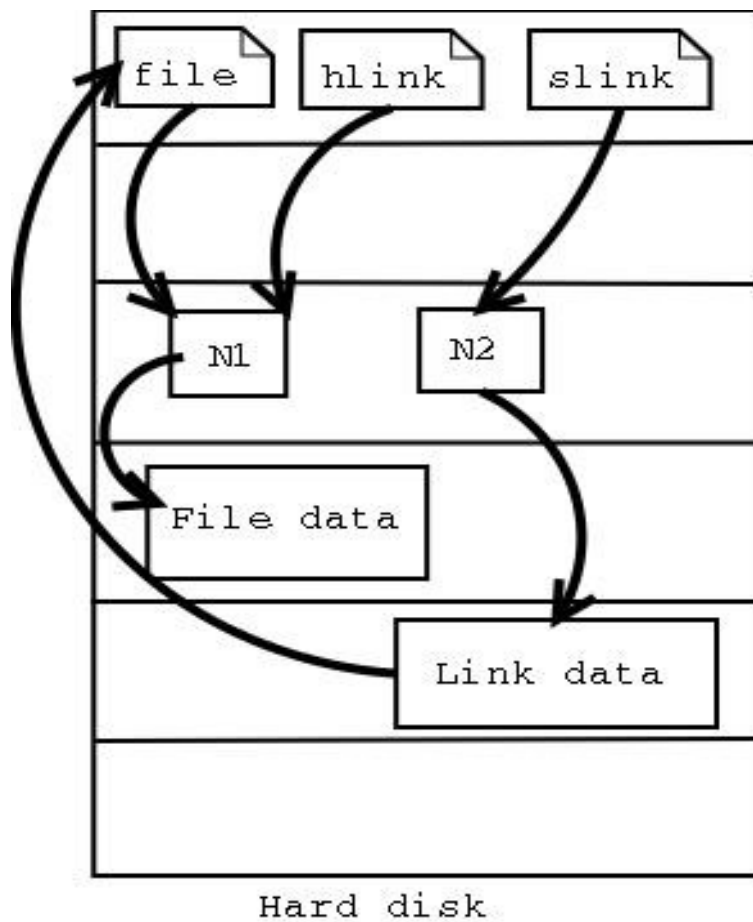
Αρχεία που δείχνουν σε άλλα αρχεία.

Οικονομία χώρου,
Καλύτερη οργάνωση,
Ευελιξία σε τροποποιήσεις

Δύο τύποι: Hard – Soft Links

Φέρονται παρόμοια αλλά υλοποιούνται
διαφορετικά

Τύποι συνδέσμων αρχείων



Δημιουργία συνδέσμων αρχείων

```
freddy:~/music> ln -s /opt/mp3/Queen/ Queen

freddy:~/music> ls -l
lrwxrwxrwx  1 freddy  freddy  17 Jan 22 11:07 Queen ->
/opt/mp3/Queen
```

Με τη εντολή

ln filename linkname

δημιουργούμε hard link, ενώ με την επιλογή -s

δημιουργούμε soft link.

Απλή αναζήτηση αρχείων: which

Χρήση χαρακτηριστικών Bash: συμπλήρωση ονομάτων διαδρομής με διαδοχικά Tab.

Εντολή which: δίνει το όνομα διαδρομής για αρχεία που βρίσκονται σε καταλόγους του PATH, π.χ.

```
which acroread
```

Η επιλογή -a ελέγχει για συνώνυμα, π.χ.

```
which -a ls
```

Απλή αναζήτηση αρχείων: file

Εντολή file: εμφανίζει ιδιότητες αρχείου, πχ

```
kmarg@pdptest1:~$ file hello.c
hello.c: ASCII C program text
kmarg@pdptest1:~$ file hier
hier: UTF-8 Unicode English text
kmarg@pdptest1:~$ file temp
temp: directory
```

Απλή αναζήτηση αρχείων: find

Απλή χρήση της εντολής find:

```
find [directory_path] -name filename
```

Παράδειγμα:

```
kmarg@pdptest1:/$ sudo find /home -name mozilla
/home/kmarg/.gnome2/devhelp/mozilla
kmarg@pdptest1:/$ sudo find / -name mozilla
/var/lib/dpkg/alternatives/mozilla
/var/lib/mozilla
/home/kmarg/.gnome2/devhelp/mozilla
/etc/alternatives/mozilla
/etc/mozilla
/usr/bin/mozilla
/usr/lib/mozilla
/usr/share/ca-certificates/mozilla
/usr/share/texmf/tex4ht/ht-fonts/mozilla
```

Απλή αναζήτηση αρχείων: locate

Απλή χρήση της εντολής locate:

`locate [options] [directory_path] string`

Παράδειγμα:

```
kmarg@pdptest1:~$ locate bashrc
/home/kmarg/.bashrc
/home/pdpuser/.bashrc
/home/admin/.bashrc
/etc/skel/.bashrc
/etc/bash.bashrc
/root/.bashrc
/usr/share/base-files/dot.bashrc
/usr/share/doc/adduser/examples/adduser.local.conf.examples/skel/dot.bashrc
/usr/share/doc/adduser/examples/adduser.local.conf.examples/bash.bashrc
```

Οικογένεια grep

Για αναζήτηση ενός προτύπου ή αλφαριθμητικού μέσα σε αρχεία κειμένου χρησιμοποιούμε την εντολή grep και τις παραλλαγές της (fgrep και egrep).

Οι εντολές grep, fgrep και egrep εμφανίζουν όλες τις γραμμές του αρχείου κειμένου που περιέχουν το πρότυπο.

Οι εντολές αυτές είναι παρόμοιες αλλά κάθε μια από αυτές έχουν ελαφρά διαφορετική συμπεριφορά.

Οικογένεια grep

Σύνταξη: `grep [options] pattern filename(s)`

`fgrep [options] string filename(s)`

`egrep [options] pattern filename(s)`

Αν δεν οριστούν τα αρχεία, τότε κάνει αναζήτηση στην πρότυπη είσοδο (πληκτρολόγιο).

Σημαντικές επιλογές εντολών grep:

n: εμφανίζει τους αριθμούς γραμμών.

i: αγνοεί τη διάκριση μεταξύ πεζών και κεφαλαίων χαρακτήρων.

l: εμφανίζει τα ονόματα αρχείων που περιέχουν το πρότυπο.

v: εμφανίζει τις γραμμές κειμένου που δεν ταιριάζουν το πρότυπο.

w: περιορίζει την αναζήτηση σε ολόκληρες λέξεις μόνο.

Οικογένεια grep: Παραδείγματα

```
cp /etc/passwd passwd1
```

```
grep rdpuser passwd1
```

Εμφανίζει όλες τις γραμμές του αρχείου passwd1 που περιέχουν το πρότυπο rdpuser.

```
grep -wn bash passwd1
```

Εμφανίζει όλες τις γραμμές του αρχείου passwd1 αριθμημένες που περιέχουν το πρότυπο bash σε ολόκληρη λέξη μόνο.

Οικογένεια grep

Οι εντολές grep, fgrep και egrep διαφέρουν μεταξύ τους κατά ότι το καθένα επιτρέπει να ταυτίζει ένα διαφορετικό είδος προτύπου.

Η εντολή grep χρησιμοποιείται για αναζήτηση μιας κανονικής έκφρασης.

Η εντολή egrep χρησιμοποιείται για αναζήτηση μιας εκτεταμένης κανονικής έκφρασης.

Η εντολή fgrep χρησιμοποιείται για αναζήτηση ενός σταθερού string.

Κανονικές εκφράσεις (1)

Κανονική έκφραση (regular expression) είναι ένα πρότυπο αποτελούμενο από μια ακολουθία χαρακτήρων που ταυτίζεται μέσα στο κείμενο.

Οι κανονικές εκφράσεις χρησιμοποιούνται:

- Για να προσδιορίσουν ονόματα αρχείων στο σύστημα αρχείων.

- Για αναζήτηση αλφαριθμητικών στα αρχεία κειμένου.

- Για εξαγωγή δεδομένων από μια βάση δεδομένων.

Οι κανονικές εκφράσεις προτάθηκαν για πρώτη φορά το 1956 (S. Kleene).

Ο Ken Thomson (1968) τις χρησιμοποίησε στην εντολή grep (global regular expression print).

Κανονικές εκφράσεις (2)

Προγράμματα που υποστηρίζουν κανονικές εκφράσεις
είναι:

grep

egrep

vi, vim, emacs

sed

awk, nawk

perl

Κανονικοί χαρακτήρες και μετα-χαρακτήρες

Οι κανονικές εκφράσεις αποτελούνται από κανονικούς χαρακτήρες και μετα-χαρακτήρες.

Οι κανονικοί χαρακτήρες είναι ένα οποιοδήποτε χαρακτήρα εκτός τον χαρακτήρα νέας γραμμής.

Οι μετα-χαρακτήρες είναι ειδικοί χαρακτήρες που έχουν ειδική σημασία.

Μετα-χαρακτήρες (1)

Η τελεία (.) ταυτίζει έναν οποιοδήποτε απλό χαρακτήρα εκτός από τον χαρακτήρα νέας γραμμής.

α.c ταυτίζει abc, adc, a&c και a;c.

u..x ταυτίζει unix, unax και u3(x.

Ο αστερίσκος (*) ταυτίζει μηδέν ή περισσότερες εμφανίσεις του χαρακτήρα που προηγείται.

ab*c ταυτίζει ac, abc, abbc και abbbc.

.* ταυτίζει οποιοδήποτε αλφαριθμητικό.

Μετα-χαρακτήρες (2)

Κλάση χαρακτήρων ([]) ταυτίζει οποιοδήποτε από τους χαρακτήρες που βρίσκονται μέσα στις αγκύλες.

Η παύλα (-) δηλώνει μια περιοχή χαρακτήρων. Π.χ. [a-e]

Ο χαρακτήρας (^) ταιριάζει κάθε χαρακτήρα που δεν περικλείεται μέσα στις αγκύλες.

Παραδείγματα:

[fF]un ταυτίζει fun και Fun.

b[aeiou]g ταυτίζει bag, beg, big, bog και bug.

[A-Z].* ταυτίζει ένα αλφαριθμητικό που αρχίζει με ένα κεφαλαίο γράμμα.

[^abc].* ταυτίζει ένα οποιοδήποτε αλφαριθμητικό που να μην αρχίζει με τους χαρακτήρες a, b ή c.

Μετα-χαρακτήρες (3)

Το σύμβολο (^) ταυτίζει μόνο την αρχή της γραμμής.

^D.* ταυτίζει μια γραμμή που στην αρχή της ξεκινά με D.

Σύμβολο δολάριο (\$) ταυτίζει μόνο το τέλος της γραμμής.

.*d\$ ταυτίζει μια γραμμή που τελειώνει με d.

Σύμβολο (\) αναστέλει την σημασία κάθε μετα-χαρακτήρα.

file\\.txt ταυτίζει file.txt αλλά όχι file_txt.

Μετα-χαρακτήρες και χαρακτήρες μπαλαντέρ

Για την ομαλή λειτουργία οι κανονικές εκφράσεις πρέπει να περικλείονται σε μονές αποστρόφους (π.χ. ' ') ώστε ο φλοιός να διερμηνεύσει τους χαρακτήρες που περιέχονται στην κανονική έκφραση ως μετα-χαρακτήρες και όχι ως χαρακτήρες μπαλαντέρ.

Οι χαρακτήρες μπαλαντέρ είναι ένα εργαλείο που επιτρέπει την αναφορά σε ονόματα αρχείων με συγκεκριμένα χαρακτηριστικά.

Οι μετα-χαρακτήρες που χρησιμοποιούνται στις κανονικές εκφράσεις είναι διαφορετικά και πιο ισχυρά εργαλεία αναζήτησης προτύπων. Οι μετα-χαρακτήρες χρησιμοποιούνται από ειδικές εντολές (π.χ. `grep`, `sed`).

Κανονικές εκφράσεις: Παραδείγματα 1

`grep '.n' textfile`

Εμφανίζει όλες τις γραμμές του αρχείου textfile που ταυτίζουν ένα οποιοδήποτε χαρακτήρα ακολουθούμενο από τον χαρακτήρα n.

`grep 'bio*' textfile`

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν το πρότυπο bio ακολουθούμενο από μια οποιαδήποτε ακολουθία χαρακτήρων.

`grep '199[1-5]' textfile`

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν τα έτη 1991 έως 1995.

Κανονικές εκφράσεις: Παραδείγματα 2

```
grep '^[a-z]' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που ξεκινούν με έναν οποιοδήποτε πεζό αγγλικό χαρακτήρα.

```
grep '22$' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που τελειώνουν σε 22.

```
grep '$1\.20' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν την τιμή \$1.20.

Κανονικές εκφράσεις: Παραδείγματα 3

```
grep '^ [A-Z]' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που ξεκινούν με ένα κεφαλαίο αγγλικό χαρακτήρα.

```
grep '^..$' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν 2 χαρακτήρες.

```
grep '^[^0-9]*$' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που δεν περιέχουν αριθμούς.

Κανονικές εκφράσεις: Παραδείγματα 4

```
grep '[0-9]*[.][0-9]*' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν πραγματικούς αριθμούς.

```
grep '1*[012]*[1-9]*:[0-5][0-9]' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν ώρα σε 12ωρη μορφή.

Εκτεταμένες κανονικές εκφράσεις:

Εντολή egrep

Οι εκτεταμένες κανονικές εκφράσεις υποστηρίζονται από την εντολή egrep.

Είναι υπερσύνολο της εντολής grep.

Αναζητεί και αντιστοιχεί όλες τις περιπτώσεις κανονικών εκφράσεων της grep καθώς και επιπλέον μετα-χαρακτήρων.

Μετα-χαρακτήρες egrep (1)

Το σύμβολο (+) ταυτίζει μια ή περισσότερες εμφανίσεις του χαρακτήρα που προηγείται.

$ab+c$ ταυτίζει abc , $abbc$, $abbbc$ άλλα όχι ac .

Το ερωτηματικό (?) ταυτίζει μηδέν ή μια εμφάνιση του χαρακτήρα που προηγείται.

$ab?c$ ταυτίζει ac ή abc αλλά όχι $abbc$.

Το λογικό Η' (|) ταυτίζει μια από τις δύο κανονικές εκφράσεις.

$abc|def$ ταυτίζει abc ή def .

Μετα-χαρακτήρες egrep (2)

Οι παρενθέσεις () τοποθετούνται στην κανονική έκφραση ώστε οι μεταχαρακτήρες *, + ή ?, να δράσουν στο σύνολο της έκφρασης αντί μόνο σε χαρακτήρα.

$a(bc)^*$ ταυτίζει a , abc , $abcbc$, $abcbcbc$.

$(foot|base)ball$ ταυτίζει $football$ ή $baseball$.

Οι αγκύλες { } δηλώνουν τον αριθμό των επαναλήψεων που μπορεί να επαναληφθεί η κανονική έκφραση.

$[a-z]\{3\}$ ταιριάζει τρεις πεζούς χαρακτήρες.

$m.\{2,4\}$ ταιριάζει αλφαριθμητικά που ξεκινούν με m και ακολουθούν από 2 μέχρι και 4 χαρακτήρες.

Εκτεταμένες κανονικές εκφράσεις:

Παραδείγματα 1

```
egrep '[A-Z][A-Z]+' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν μια οποιαδήποτε ακολουθία μόνο κεφαλαίων αγγλικών χαρακτήρων.

```
egrep '[α-ζ]?' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν ένα ή κανένα πεζό αγγλικό χαρακτήρα.

```
egrep '239(4|6)-(0|1)' textfile
```

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν τηλέφωνα που το 4 ψηφίο είναι 4 ή 6 και το 6 ψηφίο είναι 0 ή 1.

Εκτεταμένες κανονικές εκφράσεις:

Παραδείγματα 2

`egrep '[0-9]{2}' textfile`

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν αριθμούς με 2 ψηφία.

`egrep '[\!*]' textfile`

Εμφανίζει όλες τις γραμμές του αρχείου textfile που περιέχουν τους χαρακτήρες ! και *.

Εντολή fgrep

Χρησιμοποιείται για αναζήτηση σταθερών αλφαριθμητικών (strings) και δεν χρησιμοποιεί μετα-χαρακτήρες.

Στην εντολή fgrep υπάρχει η επιλογή -f pattern_filename για να αναζητήσουμε πρότυπα που βρίσκονται στο αρχείο προτύπων με όνομα pattern_filename.

```
fgrep -f pattern_filename filename
```

pattern_filename περιέχει τα πρότυπα.

Εντολή fgrep: Παραδείγματα

`fgrep friend datafile`

Εμφανίζει όλες τις γραμμές του αρχείου `datafile` που περιέχουν το σταθερό αλφαριθμητικό `friend`.

`fgrep -f pattern.txt textfile`

Εμφανίζει όλες τις γραμμές του αρχείου `textfile` που περιέχουν τα πρότυπα που περιλαμβάνονται στο αρχείο `pattern.txt`.

Σύνθετη αναζήτηση αρχείων (1)

Το εργαλείο `find` είναι το πλέον διαδεδομένο και δυνατό εργαλείο για σύνθετη αναζήτηση αρχείων διότι κάνει πολύ περισσότερα πράγματα από απλά να βρίσκει αρχεία.

Το `find` μπορεί να εκτελέσει ενέργειες σε ένα σύνολο αρχείων που ικανοποιούν συγκεκριμένα κριτήρια, ακόμη και συνδυασμό κριτηρίων.

Οι ενέργειες της `find` περιλαμβάνουν τη διαγραφή, αντιγραφή, μετονομασία αρχείων κλπ.

Σύνθετη αναζήτηση αρχείων (2)

Τα κριτήρια αναζήτησης της find μπορεί να είναι:

Αναζήτηση με το χρόνο προσπέλασης ή τύπο ή μέγεθος του αρχείου.

Αναζήτηση με το όνομα ή με κάποια κανονική έκφραση.

Παράδειγμα:

Πρέπει να διαγράψουμε όλα τα αρχεία τα οποία έχουν μέγεθος πάνω από 50 Kbytes και δεν έχουν τροποποιηθεί για 3 μέρες.

Εντολή find: γενική σύνταξη

Σύνταξη: `find pathname expression action`

Το `pathname` είναι το σημείο εκκίνησης της αναζήτησης.

Το `expression` είναι μια έκφραση που συνδυάζει ένα ή περισσότερα κριτήρια αναζήτησης.

Η `action` είναι μια ενέργεια που θα επιδράσει πάνω στα αρχεία που έχουν εντοπισθεί.

Εντολή find: επιλογές (1)

Εκφράσεις της εντολής find:

- name filename: αληθής αν το όνομα αρχείου ταυτίζεται με το filename.
- type filetype: αληθής αν το τύπος του αρχείου είναι
 - b – αρχείο μπλοκ
 - c - αρχείο χαρακτήρα
 - d - κατάλογος
 - f – κανονικό αρχείο
 - p – ονομαστικός διάυλος fifo
 - l – συμβολικός σύνδεσμος
 - s - υποδοχή

Εντολή find: επιλογές (2)

Εκφράσεις της εντολής find:

- size +/-number: αληθής αν το μέγεθος του αρχείου είναι number.

- atime +/- n: αληθής αν το αρχείο έχει προσπελαστεί μέσα n μέρες.

Το κενό ανάμεσα στα κριτήρια είναι το λογικό ΚΑΙ. Το -ο ανάμεσα στα κριτήρια είναι το λογικό Η. Το ! είναι το λογικό ΌΧΙ.

Εντολή find: επιλογές (3)

Ενέργειες της εντολής find:

- print: εμφανίζει όλα τα αρχεία που έχουν εντοπισθεί, στην οθόνη του υπολογιστή.
- exec command \; :εκτελεί την εντολή command στα αρχεία που έχουν εντοπισθεί. Αν καθορίσετε {} σαν όρισμα γραμμής εντολών, αυτό αντικαθίσταται από το όνομα του τρέχοντος αρχείου.
- ok command \; : Το ίδιο με το exec αλλά με ερώτηση επιβεβαίωσης του χρήστη πριν από την εκτέλεση της εντολής command.

Εντολή find: Παραδείγματα 1

```
find / -type d -print
```

Αναζητά όλους τους καταλόγους σε όλο το σύστημα Linux και εμφανίζει τα ονόματα των καταλόγων.

```
find / -type d -size 30 -print
```

Αναζητά όλους τους καταλόγους που έχουν μέγεθος 30 blocks σε όλο το σύστημα Linux και εμφανίζει τα ονόματα των καταλόγων.

```
find /etc -name 'f*' -print
```

Αναζητά ονόματα αρχείων που ξεκινούν με f από το κατάλογο /etc και εμφανίζει τα ονόματα των αρχείων.

Εντολή find: Παραδείγματα 2

```
find / -name core -exec rm -f '{}'\;
```

Αναζητά αρχεία με όνομα core σε όλο το σύστημα και όσα εντοπιστούν να διαγραφούν.

```
find / -size +100k -exec cp '{}' bigfiles \;
```

Αναζητά αρχεία που έχουν μέγεθος μεγαλύτερο από 100 KB σε όλο το σύστημα και όσα εντοπιστούν να μετακινηθούν στον κατάλογο bigfiles.

```
find . -name '*.c' -o -name '*.txt' -print
```

Αναζητά αρχεία που τελειώνουν σε c ή σε txt από το τρέχοντα κατάλογο και εμφανίζει τα ονόματά τους.