



Πανεπιστήμιο Δυτικής Μακεδονίας
Τμήμα Μηχανικών Πληροφορικής & Τηλεπικοινωνιών

Λειτουργικά Συστήματα

Ενότητα 10: Διαχείριση Αρχείων

Δρ. Μηνάς Δασυγένης

mdasyg@ieee.org

Εργαστήριο Ψηφιακών Συστημάτων και Αρχιτεκτονικής Υπολογιστών

<http://arch.ict.e.uowm.gr/mdasyg>

Τμήμα Μηχανικών Πληροφορικής και Τηλεπικοινωνιών



Πανεπιστήμιο Δυτικής Μακεδονίας



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ψηφιακά Μαθήματα στο Πανεπιστήμιο Δυτικής Μακεδονίας**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
Πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ



Σκοπός Ενότητας

- Η κατανόηση της σημαντικότητας των συστημάτων αρχείων.
- Η περιγραφή των αρχιτεκτονικών των συστημάτων αρχείων.
- Η παρουσίαση τυπικών συστημάτων αρχείων δημοφιλών ΛΣ.



Υπάρχει η ανάγκη της μακροχρόνιας αποθήκευσης πληροφοριών

- Δυνατότητα αποθήκευσης **μεγάλων** ποσοτήτων πληροφοριών (μερικές φορές και μεγαλύτερου από το μέγεθος του ιδεατού χώρου διευθύνσεων).
- Οι πληροφορίες πρέπει να **επιβιώνουν** μετά τον τερματισμό της διεργασίας που τις χρησιμοποιεί (διατήρηση δεδομένων).
- Πρέπει να επιτρέπεται σε πολλές διεργασίες να έχουν το δικαίωμα **ταυτόχρονης προσπέλασης** στις πληροφορίες (π.χ. πρόσβαση σε μια βάση δεδομένων).



Μια απλή υλοποίηση μακροχρόνια αποθήκευσης δε λύνει όλα τα προβλήματα

- Θα μπορούσε να χρησιμοποιηθεί ένα μέσο αποθήκευσης που να υποστηρίζει μόνο δυο λειτουργίες:
 - Ανάγνωση του k block.
 - Εγγραφή του k block.
- Ανακύπτουν πολλά προβλήματα όπως:
 - Πως θα βρούμε τις πληροφορίες;
 - Πως θα διατηρηθεί η ασφάλεια των δεδομένων;
 - Πως θα ξέρουμε ποια block χρησιμοποιούνται και ποια όχι;
- Το ΛΣ έχει απλοποιήσει αυτά τα θέματα με την χρήση του συστήματος αρχείων.



Η ανάγκη λύνεται με τη χρήση των αρχείων

- Τα αρχεία είναι ένας αφαιρετικός μηχανισμός διαχείρισης των δεδομένων των διεργασιών.
- Κάθε αρχείο είναι μια μονάδα του λειτουργικού συστήματος που βρίσκεται σε μέσα αποθήκευσης (σκληρό δίσκο, cd, dvd, USB flash disk κ.α.).
- Έχει αρκετές χρήσιμες ιδιότητες:
 - Γραμμικός αποθηκευτικός χώρος.
 - Ανεξαρτησία από αποθηκευτική συσκευή.
 - Ονοματολογία για εύκολη αναζήτηση.
 - Ευελιξία όσον αφορά τα δεδομένα προς αποθήκευση.
- Το τμήμα του ΛΣ που ασχολείται με τα αρχεία ονομάζεται **σύστημα αρχείων**.



Υπάρχουν ποικίλες δομές αρχείων...(1/2)

- Δεν υπάρχει.
 - ένα αρχείο είναι ένας γραμμικός αποθηκευτικός χώρος.
- Δεδομένα μπορεί να δομηθούν με πολλούς τρόπους:
 - Ξεχωριστή εγγραφή σε κάθε γραμμή.
 - Συνεχόμενες εγγραφές συγκεκριμένου μεγέθους.
 - Συνεχόμενες εγγραφές μεταβλητού μεγέθους.
 - Πολύπλοκη «γραμματική» δομής(SGML,XML).
- Δεδομένα μπορεί να είναι σε δυαδική μορφή ή ASCII.



Υπάρχουν ποικίλες δομές αρχείων...(2/2)

- Αρχεία μπορεί να έχουν ιδιαίτερη «σημασία»:
 - Εκτελέσιμα αρχεία.
 - Αρχεία δεδομένων εφαρμογών.
- Σημασία και δομή των δεδομένων ορίζεται από το πρόγραμμα που δημιουργεί το αρχείο.



Ένα αρχείο εκτός από δεδομένα έχει και Μεταδεδομένα

- **Όνομα:** περισσότερο για το χρήστη.
- **Τύπος:** για καλύτερο έλεγχο χρήσης (ο τύπος μπορεί και να ορίζεται έμμεσα, π.χ. μέσω ειδικής κατάληξης).
- **Τοποθεσία:** θέση των δεδομένων στο δίσκο.
- **Μέγεθος:** τρέχον μέγεθος του αρχείου.
- **Προστασία:** ρύθμιση πρόσβασης και χρήσης.
- **Χρονοσφραγίδα (timestamp):** ώρα δημιουργίας/τελευταίας αλλαγής/ πρόσβασης.

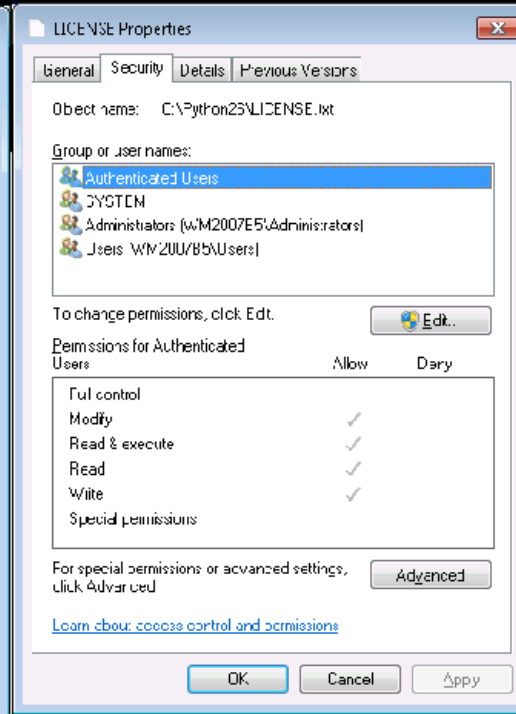
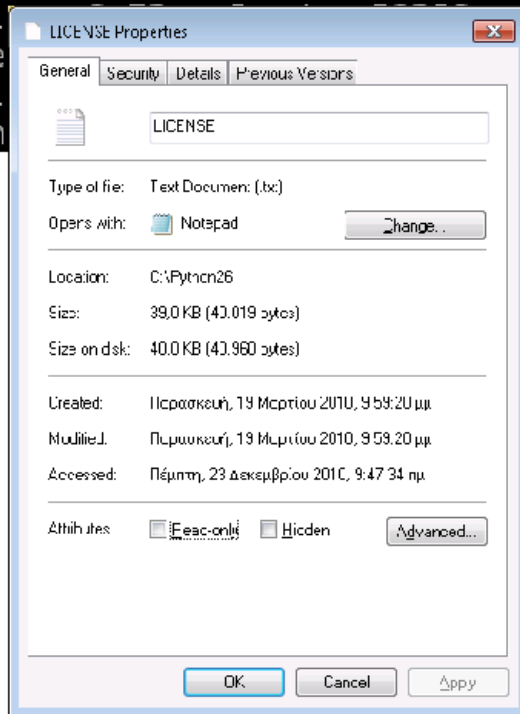


Μεταδεδομένα σε διάφορα συστήματα αρχείων...

- FreeBSD

```
mdasyg@bigb5.homeftp.net [176]: stat -x /etc/passwd
File: "/etc/passwd"
Size: 2106          FileType: Regular File
Mode: (0644/-rw-r--r--)      Uid: (   0/   root)  Gid: (   0/  wheel)
```

Devi
Acce
Modi
Chan



- Windows



Ονόματα αρχείων

- Χρησιμοποιείται για την εύρεση αρχείων αφού δημιουργηθούν.
- Κάθε αρχείο έχει ένα όνομα.
- Το όνομα αποτελείται από χαρακτήρες που είναι έγκυροι χαρακτήρες ονοματολογίας στο λειτουργικό σύστημα.
Π.χ. 124.txt, document.txt.
- Τα μικρά ή ΚΕΦΑΛΑΙΑ μπορεί να διαφοροποιούν το όνομα.
- Το όνομα μπορεί να συνδέεται με το είδος των δεδομένων του αρχείου, αλλά μπορεί και όχι π.χ. 'Command.com', 'essay.doc' αλλά και 'test1'.



Κάποιες καταλήξεις ονομάτων έχουν καθιερωθεί στην κοινότητα....

Extension	Meaning
file.bak	Backup file
file.c	C source program
file.gif	CompuServe Graphical Interchange Format image
file.hlp	Help file
file.html	World Wide Web HyperText Markup Language document
file.jpg	Still picture encoded with the JPEG standard
file.mp3	Music encoded in MPEG layer 3 audio format
file.mpg	Movie encoded with the MPEG standard
file.o	Object file (compiler output, not yet linked)
file.pdf	Portable Document Format file
file.ps	PostScript file
file.tex	Input for the TEX formatting program
file.txt	General text file
file.zip	Compressed archive



Με τα αρχεία συνδέονται κάποιες βασικές λειτουργίες

- Δημιουργία (κενού) αρχείου.
- Διαγραφή (υπάρχοντος) αρχείου.
- Αλλαγή ονόματος αρχείου.
- Περικοπή αρχείου.
- Γράψιμο/ Διάβασμα.
- Άνοιγμα: δημιουργία δομών πρόσβασης.
- Κλείσιμο: απελευθέρωση δομών πρόσβασης.
- Μετακίνηση δείκτη ανάγνωσης/εγγραφής.
- Αποθήκευση αλλαγών.



Μέθοδοι Πρόσβασης

- **Σειριακή πρόσβαση:**

- Διάβασμα και μετακίνηση προς τα εμπρός.
- Γράψιμο και μετακίνηση προς τα εμπρός.
- Επιστροφή στην αρχή του αρχείου.

- **Άμεση πρόσβαση:**

- Διάβασμα.
- Γράψιμο.
- Τοποθέτηση δείκτη σε συγκεκριμένο μέρος του αρχείου.



Σειριακή πρόσβαση



Άμεση Πρόσβαση...(1/2)

- Ένα αρχείο αποτελείται από λογικές εγγραφές σταθερού μεγέθους.
- Το αρχείο θεωρείται ως μια αριθμημένη ακολουθία από μπλοκ ή εγγραφές.
- Συνήθως χρησιμοποιούνται στις βάσεις δεδομένων για άμεση πρόσβαση στην πληροφορία (υπολογίζεται από το ερώτημα το μπλοκ στο οποίο περιέχεται η απάντηση).
- Για μεγαλύτερα σύνολα μπορούμε να χρησιμοποιούμε συναρτήσεις hash ή αρχεία δεικτών.



Άμεση Πρόσβαση...(2/2)

- Ο αριθμός του μπλοκ που παρέχεται από το χρήστη είναι σχετικός με την αρχή του αρχείου.
- Προφανώς η πραγματική διεύθυνση του μπλοκ στο δίσκο είναι διαφορετική.
- Με βάση την άμεση πρόσβαση μπορούν να δημιουργηθούν και άλλες μέθοδοι που συνήθως συνεπάγονται τη δημιουργία ενός δείκτη για το αρχείο.
- Οι δείκτες περιέχουν παραπομπές στα διάφορα μπλοκ.



Τα αρχεία ομαδοποιούνται σε καταλόγους

- Ομαδοποίηση αρχείων.
- Ένας κατάλογος μπορεί να περιέχει πολλά αρχεία.
- Για κάθε κατάλογο κρατιέται πληροφορία (όνομα, πρόσβαση, χρονοσφραγίδα, κλπ) ανάλογη με αυτή ενός αρχείου.
- Ουσιαστικά ένας κατάλογος μπορεί να θεωρηθεί σαν ειδικό αρχείο τα περιεχόμενα του οποίου είναι οι πληροφορίες που αντιστοιχούν στα αρχεία που περιέχει ο κατάλογος.



Με τους καταλόγους συνδέονται κάποιες βασικές λειτουργίες

- Δημιουργία (κενού) κατάλογου.
- Διαγραφή (υπάρχοντος) κατάλογου.
- Αλλαγή ονόματος καταλόγου.
- Πρόσβαση στα περιεχόμενα του καταλόγου.
- Δημιουργία αρχείου στον κατάλογο.
- Σβήσιμο αρχείου από τον κατάλογο.
- Αλλαγή ονόματος αρχείου του καταλόγου.



Συστήματα Καταλόγων...(1/2)

- Επίπεδο σύστημα: ένας μοναδικός κατάλογος.
 - απλό... Αλλά δημιουργεί προβλήματα ονοματολογίας και ομαδοποίησης των αρχείων.
- Ιεραρχικό σύστημα: ένας κατάλογος μπορεί να περιέχει άλλους καταλόγους.
 - Δεν υπάρχει πρόβλημα με τα ονόματα, εύκολη ομαδοποίηση, εύκολη διαχείριση και αναζήτηση, ευκολότερα κατανοητό στο χρήστη.

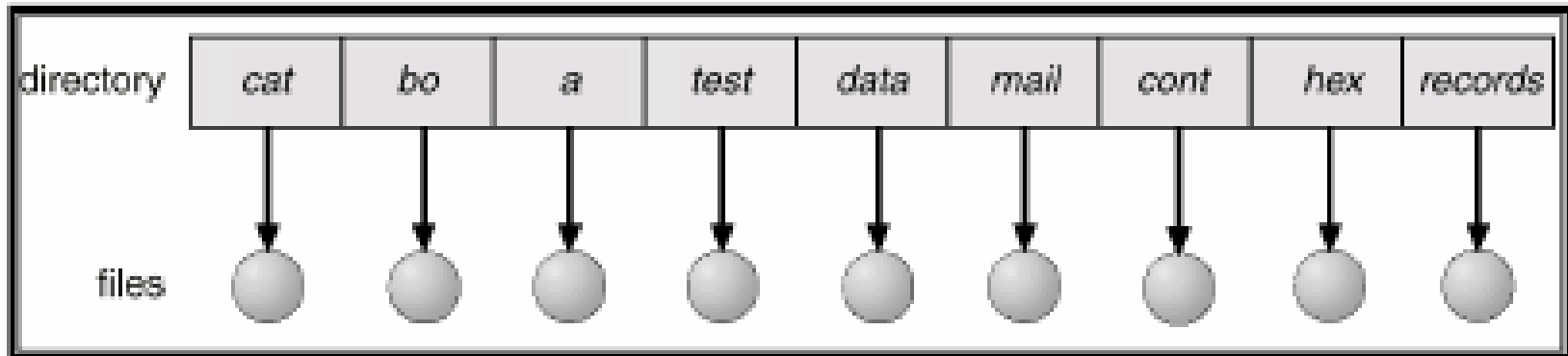


Συστήματα Καταλόγων...(2/2)

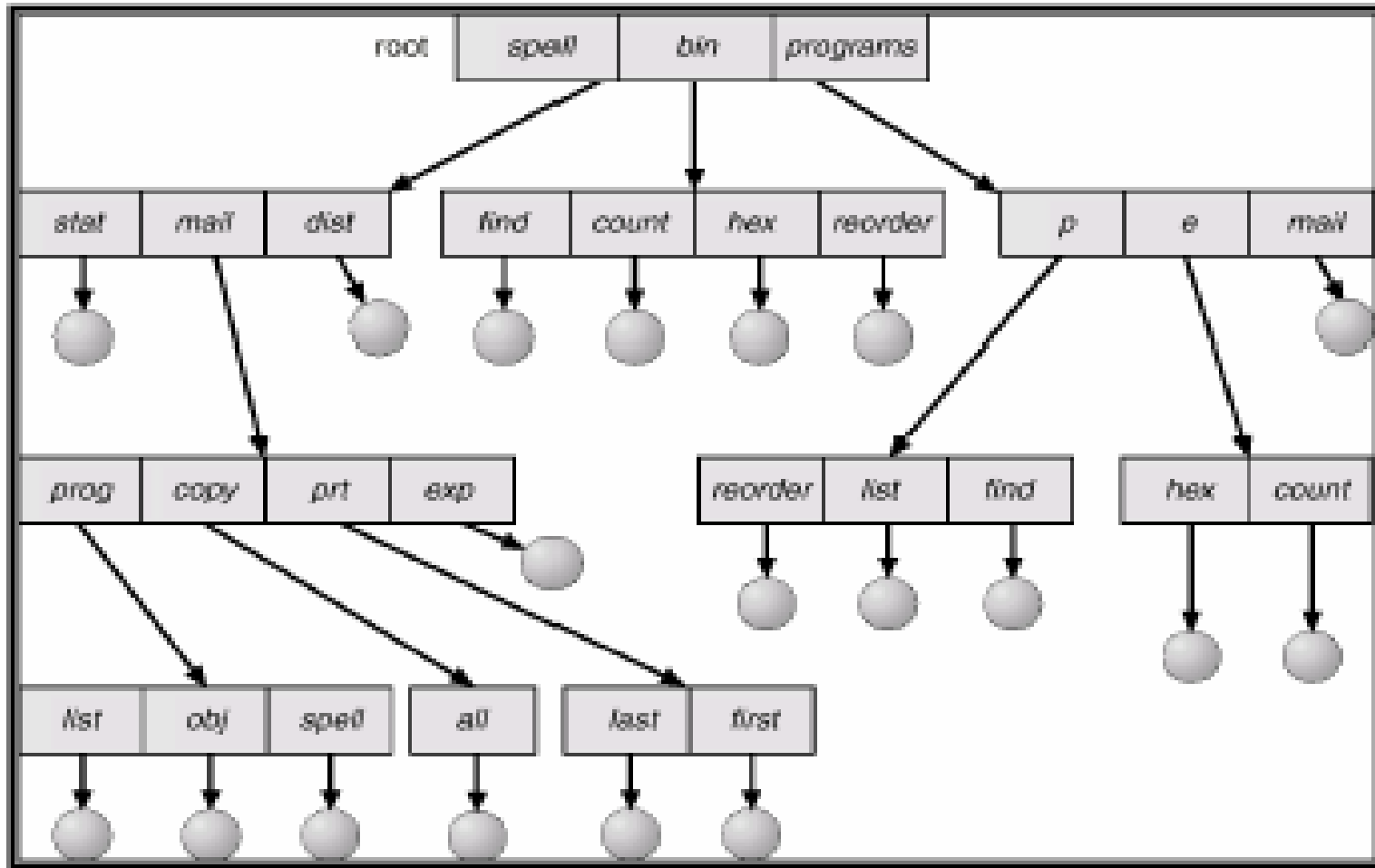
- Ιεραρχικό σύστημα με κοινά αρχεία: ένα αρχείο μπορεί να ανήκει σε πολλούς καταλόγους.
 - ευελιξία... Αλλά δημιουργεί πρόβλημα στο σβήσιμο αρχείων (λύνεται με λίστες αναφορών ή μετρητές αναφορών).
- Σύστημα με μορφή γράφου: ένας κατάλογος μπορεί να ανήκει σε πολλούς καταλόγους.
 - Αποφυγή κύκλων ή σύστημα garbage collection.



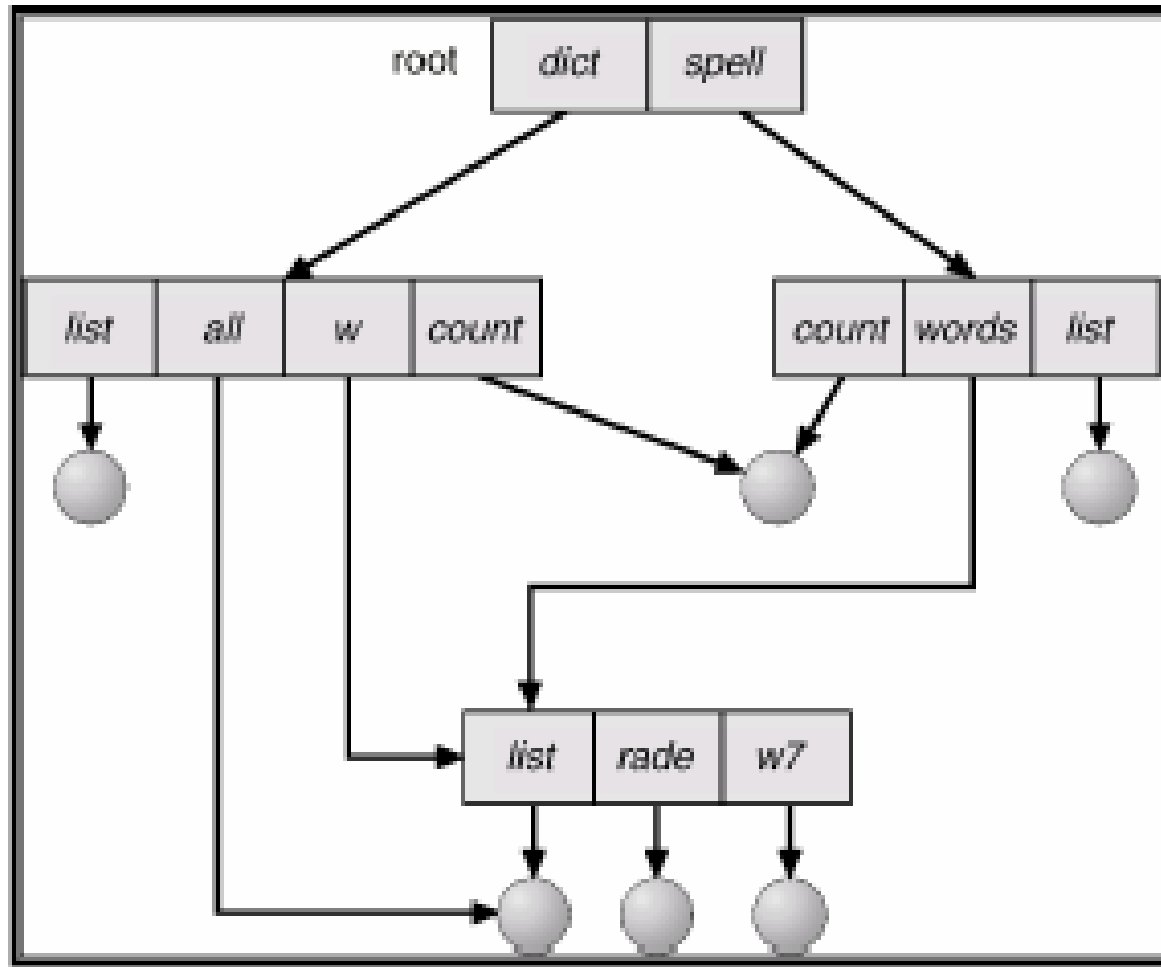
Επίπεδο σύστημα



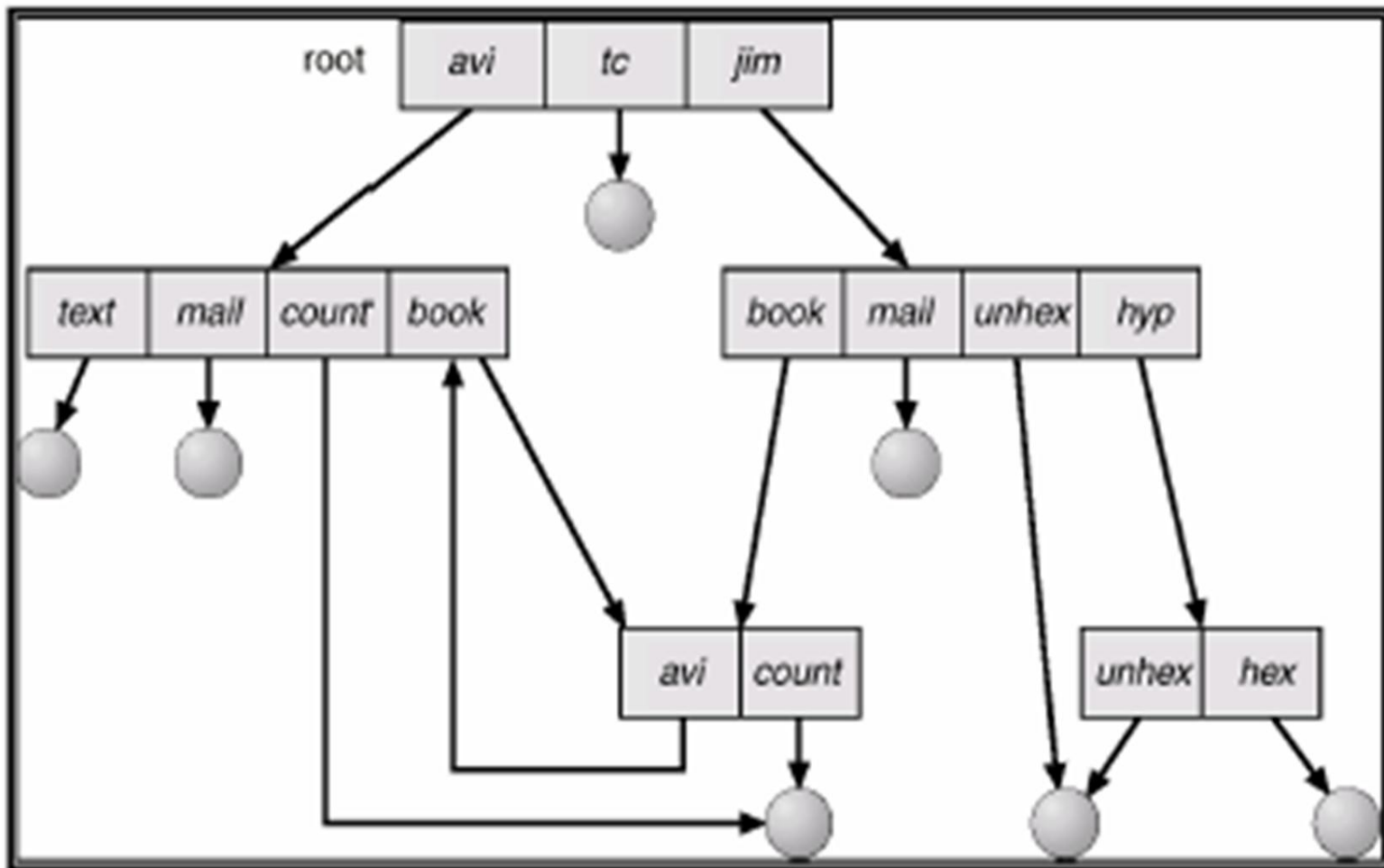
Ιεραρχικό σύστημα



Ιεραρχικό σύστημα με κοινά αρχεία



Σύστημα με μορφή γράφου



Απόλυτη & Σχετική διαδρομή (1/2)

- Σε κάθε αρχείο κατάλογο αντιστοιχεί μια απόλυτη μοναδική διαδρομή ονοματολογίας η οποία ξεκινάει από τη ρίζα.
- Κάθε αρχείο / κατάλογος έχει ένα πλήθος σχετικών διαδρομών, ανάλογα με το που βρίσκεται ο τρέχων κατάλογος. Αυτή η ονοματολογία ονομάζεται σχετική διαδρομή.
- Ειδικοί χαρακτήρες:
 - . (1 τελεία) σηματοδοτεί το τρέχον κατάλογο.
 - .. (2 τελείες) σηματοδοτεί τον προηγούμενο κατάλογο.



Απόλυτη & Σχετική διαδρομή (2/2)

- Η απόλυτη διαδρομή ξεκινά πάντα από τη ρίζα.
(unix) `/usr/local/bin/apache/htdocs/index.html`
(windows) `c:\usr\local\bin\apache\htdocs\index.html`
- Η σχετική διαδρομή ξεκινά από οποιαδήποτε θέση εκτός της ρίζας, και η διαδρομή δημιουργείται σχετικά ως προς αυτήν (ως προς το current dir):
(unix) `../htdocs/index.html`
(windows) `apache\index.html`

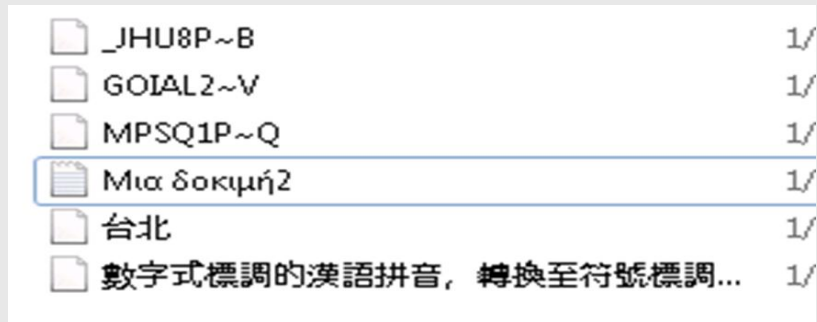


Χαρακτήρες που δεν επιτρέπονται

- (Unix) Δεν επιτρέπονται: / και null-character (\0)
Επιτρέπονται όλοι οι άλλοι:

```
total 0
-rw-r--r--  1 bigbrother  wheel  0 May  1 18:47 George<Hello>:John|
-rw-r--r--  1 bigbrother  wheel  0 May  1 18:47 Minas?
-rw-r--r--  1 bigbrother  wheel  0 May  1 18:48 Μια Απλή Δοκιμή.
-rw-r--r--  1 bigbrother  wheel  0 May  1 18:50 台北
-rw-r--r--  1 bigbrother  wheel  0 May  1 18:50 數字式標調的漢語拼音, 轉換至符號標調漢語拼音
```

- (Windows) Δεν επιτρέπονται \ / : * ? " <>



📁 _JHU8P~B	1/
📁 GOIAL2~V	1/
📁 MPSQ1P~Q	1/
📄 Μια δοκιμή2	1/
📁 台北	1/
📄 數字式標調的漢語拼音, 轉換至符號標調...	1/



Δομές Δεδομένων που τηρούνται στο δίσκο

- Boot control block:
 - Πληροφορίες που χρειάζονται για την εκκίνηση ενός λειτουργικού συστήματος.
- Volume control block:
 - Λεπτομέρειες της διαμέρισης (partition): πλήθος μπλοκ, μέγεθος μπλοκ, πλήθος και θέση ελεύθερων μπλοκ και άλλα.
- Directory structure:
 - Οργάνωση αρχείων (UFS: inodes, NTFS: master file table).
- File control block (FCB):
 - Δικαιώματα, κάτοχος, μέγεθος, θέση στο δίσκο.



Μπλοκ Ελέγχου Αρχείου (File Control Block)

Δικαιώματα αρχείου

Ημερομηνίες αρχείου (δημιουργίας, πρόσβασης, εγγραφής)

Ιδιοκτήτης αρχείου, γκρουπ, ACL

Μέγεθος αρχείου

Μπλοκ δεδομένων αρχείου ή δείκτες σε μπλοκ δεδομένων αρχείου



Προστασία αρχείου στο UNIX

✓ 3 επίπεδα πρόσβασης

- (r)ead.
- (w)rite.
- e(x)ecute.

			r	w	x
owner access	7	⇒	1	1	1
group access	6	⇒	1	1	0
all (public) access	1	⇒	0	0	1

✓ 3 κατηγορίες χρηστών:

- (o)wner.
- (g)roup.
- (a)ll.

✓ 3 bits πρόσβασης ανά κατηγορία χρήστη, που προσδιορίζονται μέσω της εντολής `chmod`.

πχ. `chmod 761 my_file`.

```
bigb5# ls -la /tmp/watch.log
-rw-r--r-- 1 bigbrother wheel 16506 Jun 19 08:02 /tmp/watch.log
bigb5# chmod go-r,u+x /tmp/watch.log ; ls -la /tmp/watch.log
-rwx----- 1 bigbrother wheel 16506 Jun 19 08:02 /tmp/watch.log
bigb5#
```



Δέσμευση μπλοκ για την αποθήκευση καταλόγων και αρχείων

Υπάρχουν 3 βασικά συστήματα:

- Συνεχόμενη αποθήκευση (contiguous allocation).
- Διασυνδεδεμένη αποθήκευση (linked allocation).
- Αποθήκευση με ευρετηρίαση (indexed allocation).



Συνεχόμενη αποθήκευση

- Κάθε αρχείο καταλαμβάνει συνεχόμενα μπλοκ στο δίσκο.
- Μοναδική πληροφορία που χρειάζεται να αποθηκεύσει κανείς για κάθε αρχείο είναι ο αριθμός του πρώτου μπλοκ και το μήκος του αρχείου.
- Άμεση πρόσβαση μπορεί να υλοποιηθεί πολύ αποδοτικά.

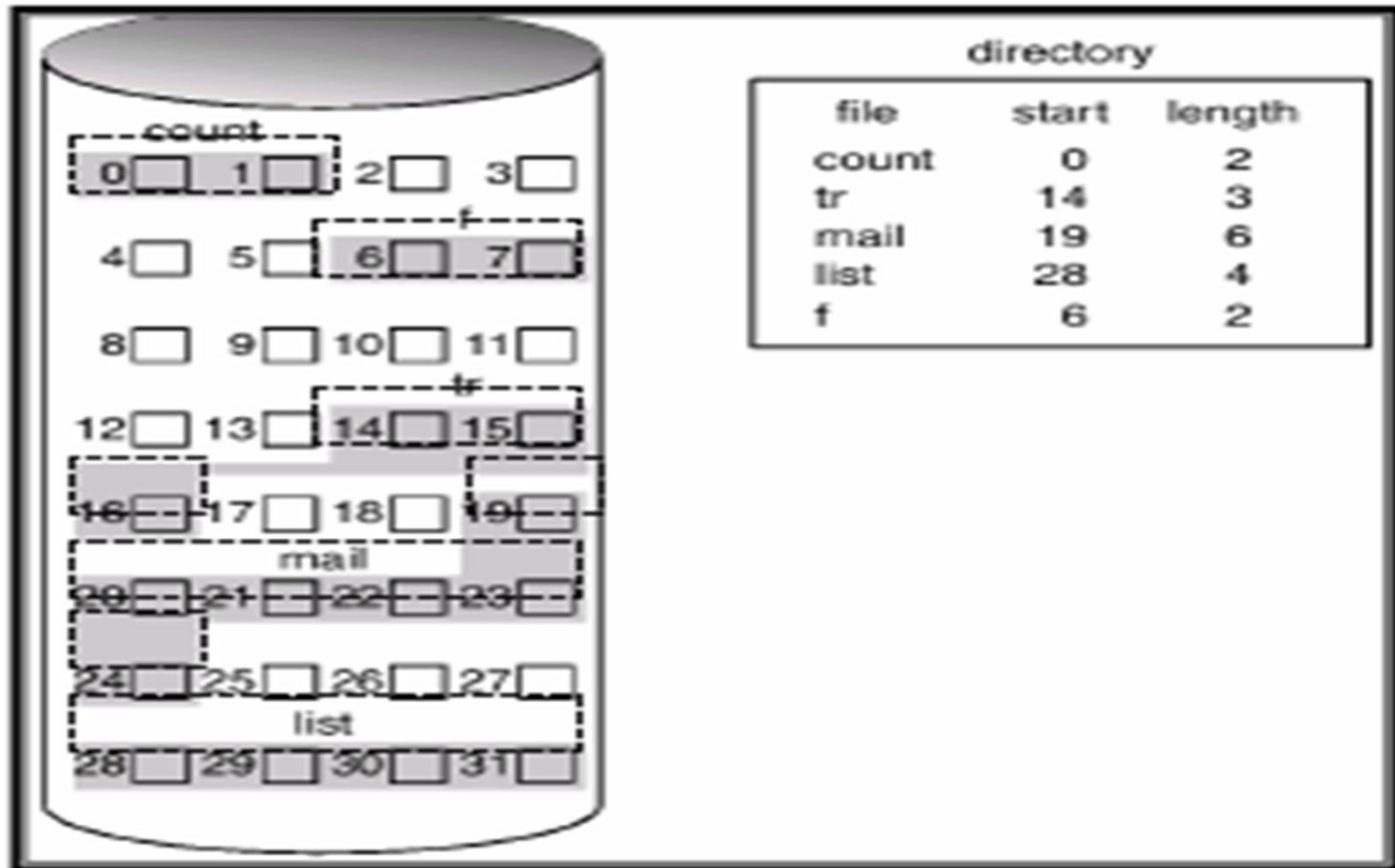


Συνεχόμενη αποθήκευση: Μειονεκτήματα

- Σπατάλη χώρου: κάποια μπλοκ που έχουν δεσμευτεί προκαταβολικά για το αρχείο παραμένουν αχρησιμοποίητα (πρόβλημα όπως και για τη δυναμική μνήμη).
- Επέκταση αρχείου: πρόβλημα αν το αρχείο μεγαλώσει και το αμέσως επόμενο μπλοκ δεν είναι ελεύθερο.



Συνεχόμενη αποθήκευση: Αναπαράσταση



Διασυνδεδεμένη αποθήκευση (1/2)

- Κάθε αρχείο αποτελείται από πολλά μπλοκ.
- Τα μπλοκ του αρχείου δεν είναι απαραίτητα συνεχόμενα.
- Κάθε μπλοκ περιέχει δείκτη στο επόμενο μπλοκ του αρχείου.
- Μοναδική πληροφορία που χρειάζεται να αποθηκευτεί για κάθε αρχείο είναι ο αριθμός του πρώτου μπλοκ του αρχείου.

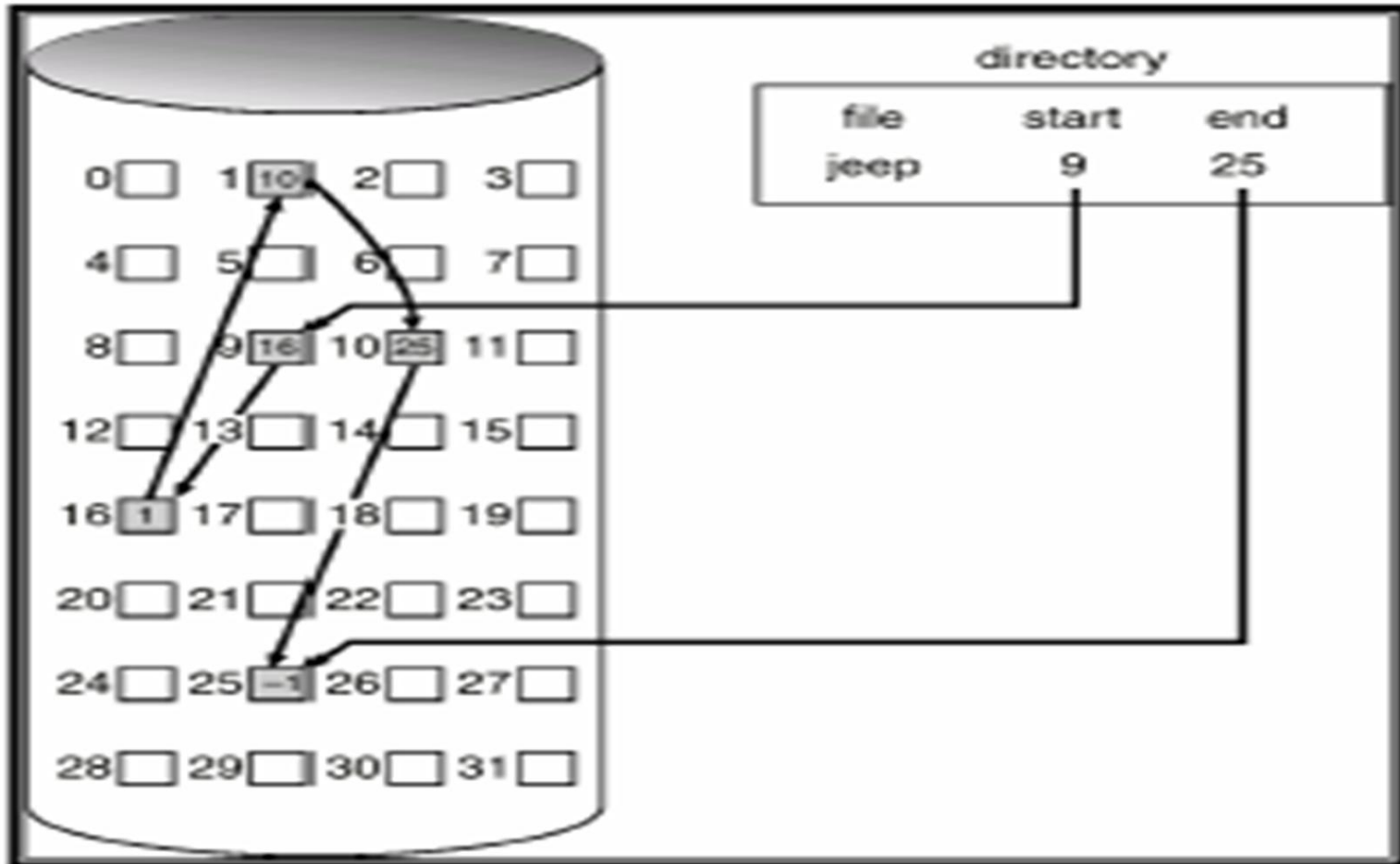


Διασυνδεδεμένη αποθήκευση (2/2)

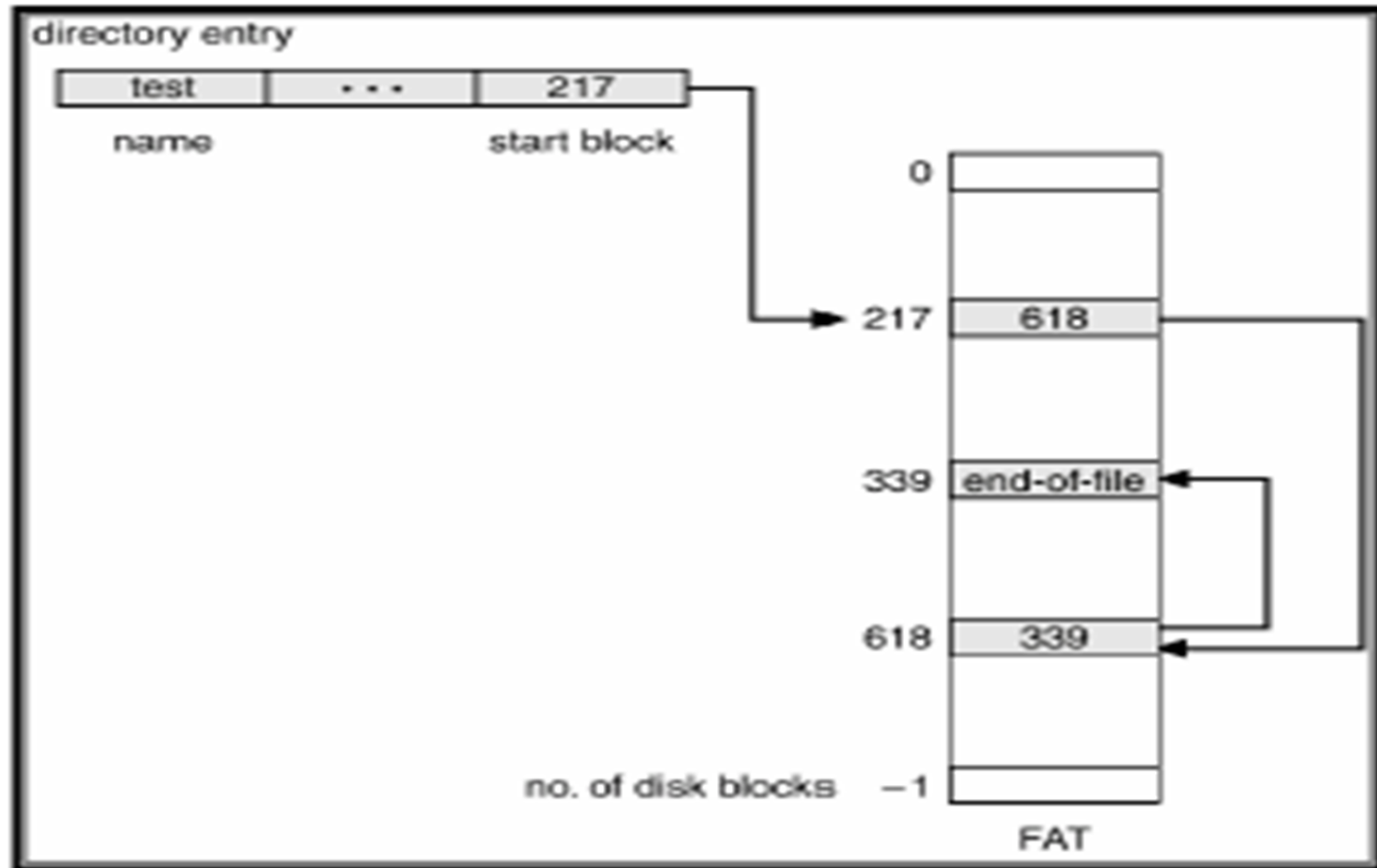
- Εύκολη επέκταση αρχείου.
- Δεν υπάρχει σπατάλη σε επίπεδο μπλοκ.
- Όμως:
 - ΔΕΝ υπάρχει φυσική υποστήριξη για άμεση πρόσβαση (μπορεί να υλοποιηθεί μέσω ειδικών δομών στη μνήμη που όμως απαιτούν να διαβαστούν τουλάχιστον μια φορά όλα τα μπλοκ του αρχείου).



Διασυνδεδεμένη αποθήκευση: Αναπαράσταση



Η δομή FAT του MS-DOS χρησιμοποιεί τη διασυνδεδεμένη αποθήκευση (1/2)



Η δομή FAT του MS-DOS χρησιμοποιεί τη διασυνδεδεμένη αποθήκευση (2/2)



Αποθήκευση με ευρετηρίαση (ή δεικτοδοτημένη ανάθεση) (1/2)

- Κάθε αρχείο αποτελείται από πολλά μπλοκ, που δεν είναι απαραίτητα συνεχόμενα.
- Διατηρείται ειδικό μπλοκ (ευρετήριο–index) που περιέχει τους αριθμούς των μπλοκ του αρχείου.
- Μοναδική πληροφορία που χρειάζεται να αποθηκευτεί για κάθε αρχείο είναι ο αριθμός του μπλοκ ευρετηρίου (index block).

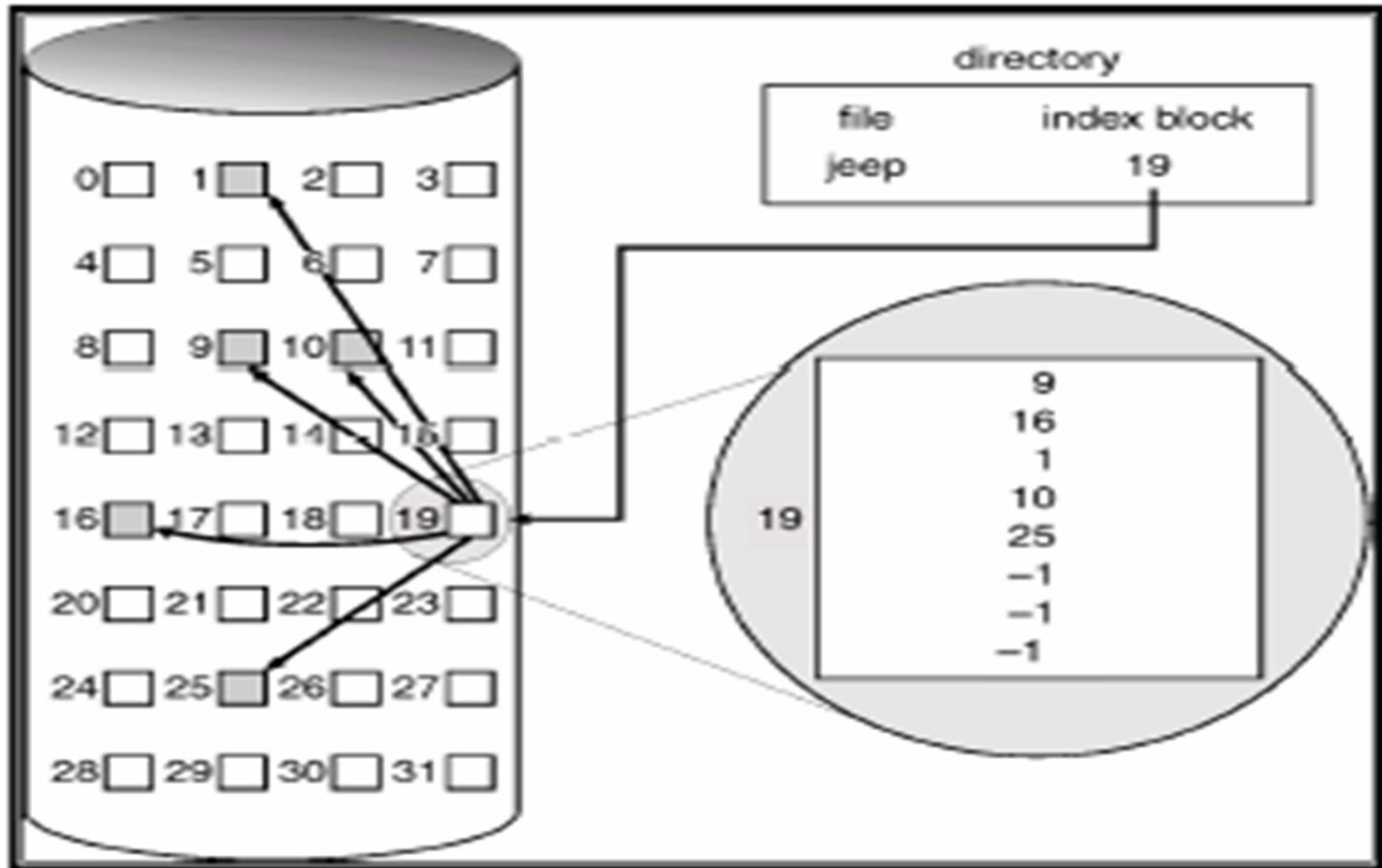


Αποθήκευση με ευρετηρίαση (ή δεικτοδοτημένη ανάθεση) (2/2)

- Εύκολη επέκταση αρχείου, και χωρίς σπατάλη σε επίπεδο μπλοκ.
- Άμεση πρόσβαση απαιτεί ανάγνωση 1 μπλοκ (που μπορεί να κρατιέται και στη μνήμη).
- Όμως:
 - αν χαθεί το ευρετήριο, χάνεται το αρχείο.



Αποθήκευση με ευρετηρίαση: Αναπαράσταση



Ευρετήρια Πολλών επιπέδων (1/2)

- Με τη μέθοδο της ευρετηρίασης το μέγεθος του αρχείου περιορίζεται από το μέγεθος του μπλοκ:
 - Μέγεθος μπλοκ 2^n bytes \rightarrow μέγιστο μέγεθος αρχείου 2^{2^n-1} bytes.
 - Για μεγαλύτερα αρχεία χρειάζεται ευρετήριο 2 επιπέδων, όπου ένα μπλοκ χρησιμοποιείται για την αποθήκευση του εξωτερικού ευρετηρίου που περιέχει τους αριθμούς των μπλοκ του ευρετηρίου.

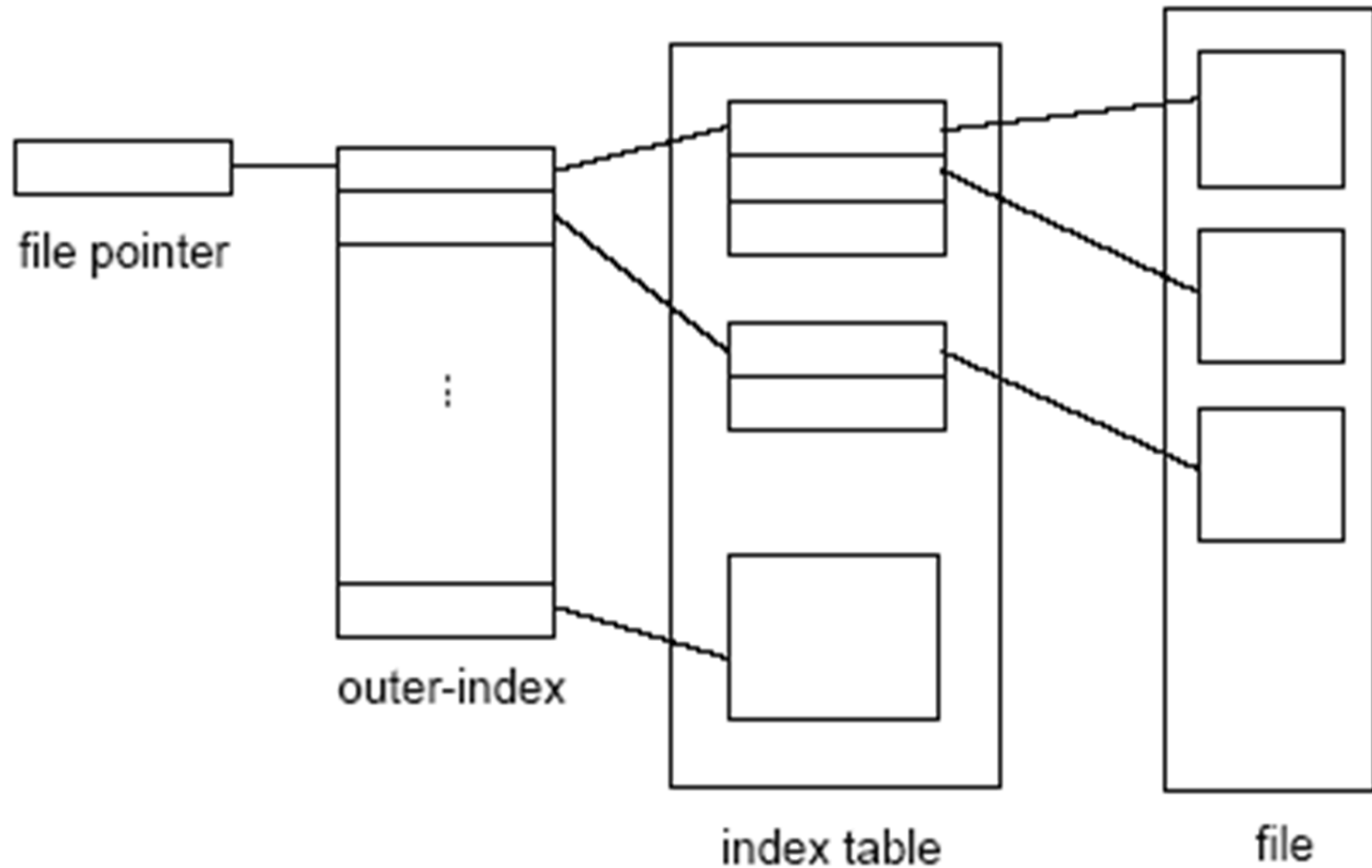


Ευρετήρια Πολλών επιπέδων (2/2)

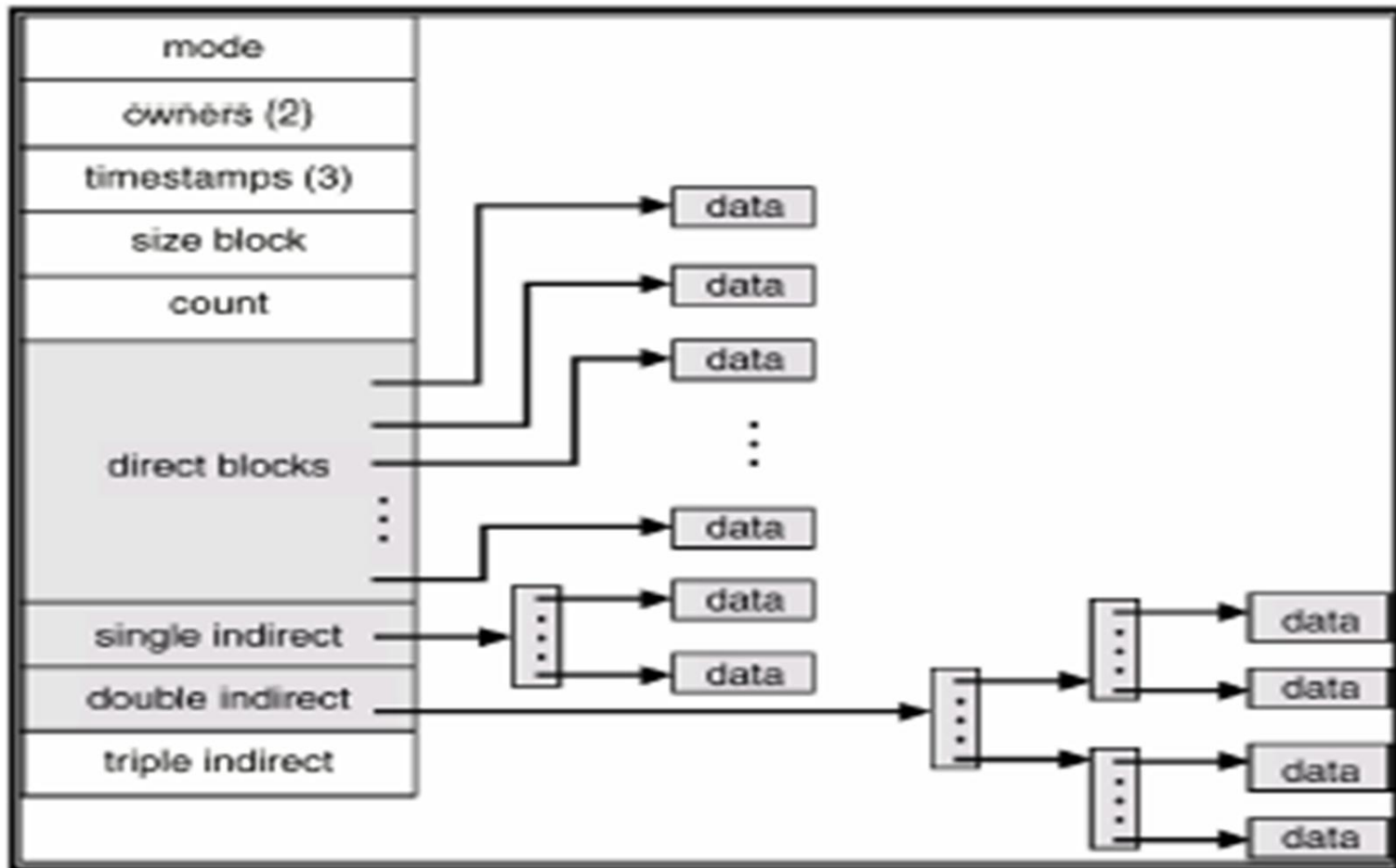
- Μοναδική πληροφορία που χρειάζεται να αποθηκευτεί για κάθε αρχείο είναι ο αριθμός του μπλοκ εξωτερικού ευρετηρίου.
- Όμως:
 - Άμεση πρόσβαση απαιτεί ανάγνωση 2 μπλοκ.
 - Αν χαθεί το εξωτερικό ευρετήριο, χάνεται το αρχείο.



Ευρετήρια 2 επιπέδων: Αναπαράσταση



Ανάθεση μπλοκ στο UNIX



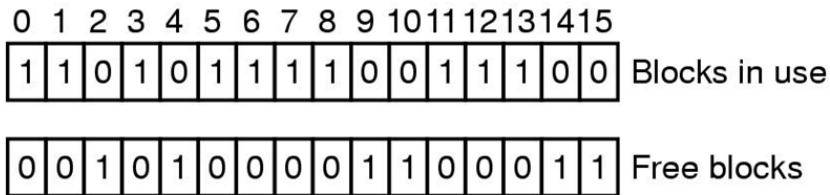
Διαχείριση ελευθέρων μπλοκ

- Χάρτης bit.
 - Κάθε μπλοκ αντιστοιχεί σε μια θέση στον πίνακα ελευθέρων μπλοκ.
 - Αν υπάρχει '1' το μπλοκ δε χρησιμοποιείται.
 - Αν υπάρχει '0' το μπλοκ είναι σε χρήση.
- Αλυσιδωτά ελεύθερα τμήματα.
 - Χρήση ενός δείκτη και μια τιμή μήκους.
- Δεικτοδότηση.
 - Αντιμετωπίζεται ως αρχείο και χρησιμοποιεί έναν πίνακα δεικτών.

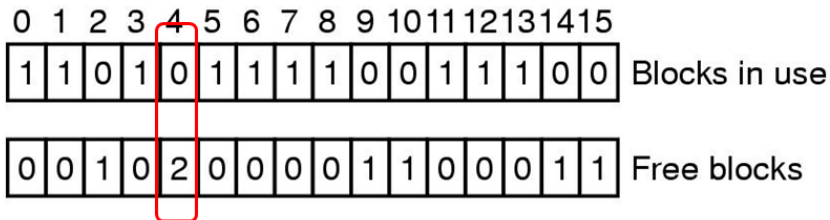


Χάρτες bit

Consistent

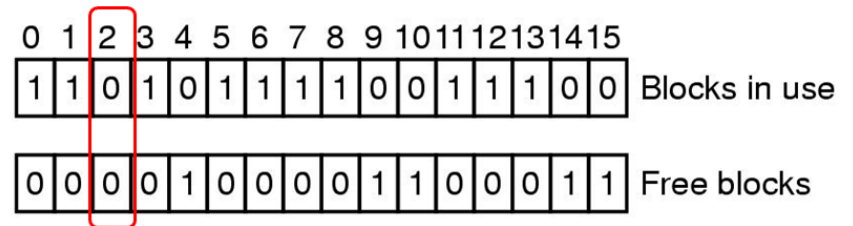


(a)

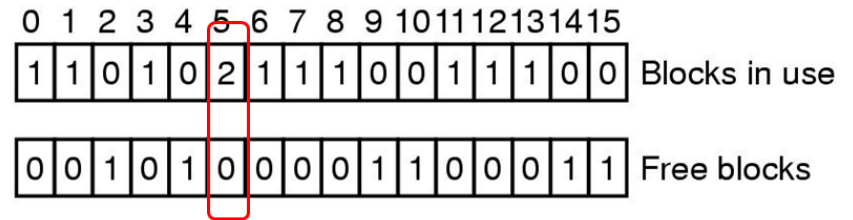


Duplicate block in free list

Missing ("lost") block



(b)



Duplicate block in two files



Άσκηση

Ένας ελεγκτής του συστήματος αρχείων έχει δημιουργήσει τους μετρητές του όπως φαίνεται στο παρακάτω σχήμα. Αυτοί είναι:

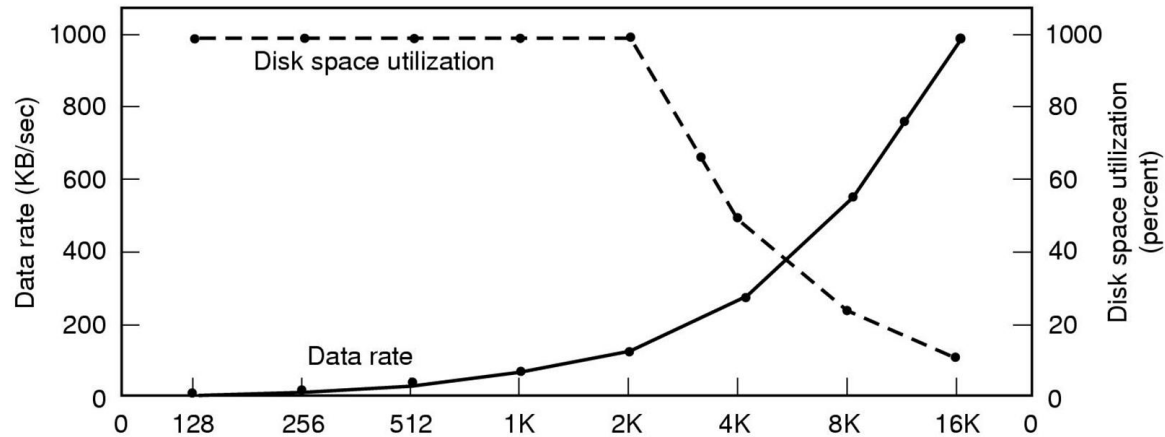
Σε χρήση: 3 1 0 1 0 2 1 0 1 1 0 1 0 0 1 0 0

Ελεύθεροι: 3 0 0 0 1 1 1 0 0 0 1 0 1 1 0 1 2

Υπάρχουν λάθη; Αν ναι, είναι σοβαρά; Γιατί; Πως θα διορθωθούν τα λάθη;



Μέγεθος εκχώρησης μπλοκ



- Ο χρόνος προσπέλασης ενός μπλοκ κυριαρχείται πλήρως από το χρόνο αναζήτησης και την καθυστέρηση λόγω περιστροφής. Όσα περισσότερα δεδομένα προσκομίζονται τόσο καλύτερη είναι η χρήση του δίσκου.
- Ο ρυθμός μεταφοράς δεδομένων αυξάνεται λοιπόν με την αύξηση του block size.
- Μειώνεται χρήση του δίσκου, γιατί αυξάνεται σημαντικά ο εσωτερικός κατακερματισμός.



Τυπικά Μεγέθη μπλοκ και μέγιστη χωρητικότητα συστήματος αρχείων

Block size	FAT-12	FAT-16	FAT-32
0.5 KB	2 MB		
1 KB	4 MB		
2 KB	8 MB	128 MB	
4 KB	16 MB	256 MB	1 TB
8 KB		512 MB	2 TB
16 KB		1024 MB	2 TB
32 KB		2048 MB	2 TB



Ανάκαμψη από λάθη...(1/2)

- Τόσο ο οδηγός του δίσκου, όσο και ένα σύστημα αρχείων μπορεί να διατηρούν κρυφές μνήμες από μπλοκ.
- Αν το σύστημα «πέσει» χωρίς να γραφτούν αλλαγές σε μπλοκ στο δίσκο, τότε χάνεται πληροφορία.
 - Περιεχόμενα αρχείων.
 - Περιεχόμενα καταλόγων.
 - Εσωτερική διαχείριση αρχείων και καταλόγων(!).



Ανάκαμψη από λάθη...(2/2)

- Ανάλογα με τη μέθοδο αποθήκευσης, μπορεί να χαθεί αρκετό μέρος από ένα αρχείο/κατάλογο.
- Αν το σύστημα αρχείων έχει υλοποιηθεί κατάλληλα, μέρος της χαμένης πληροφορίας μπορεί να ανακτηθεί διαβάζοντας όλα τα μπλοκ του δίσκου και συνδυάζοντας τα περιεχόμενα τους.



Έλεγχος Συνέπειας (1/2)

- Σε γενικές γραμμές η πληροφορία που τηρείται στην κύρια μνήμη είναι περισσότερο ενημερωμένη από την αντίστοιχη πληροφορία στο δίσκο.
- Το σύστημα αρχείων μπορεί να βρεθεί σε ασυνεπή κατάσταση (inconsistent state).
- Ένα ειδικό πρόγραμμα ελέγχει και διορθώνει ασυνέπειες.
- Συγκρίνονται τα δεδομένα στη δομή καταλόγου με τα μπλοκ δεδομένων στο δίσκο (οι αλγόριθμοι ανάθεσης και διαχείρισης ελεύθερου χώρου καθορίζουν το είδος των προβλημάτων και το εύρος της επιτυχίας στη διόρθωση).



Έλεγχος Συνέπειας (2/2)

- Στη διασυνδεδεμένη ανάθεση αν υπάρχει δείκτης από το κάθε μπλοκ στο επόμενο του, μπορεί να αναδημιουργηθεί το αρχείο και η εγγραφή του καταλόγου.
- Αντίθετα, η απώλεια μιας εγγραφής καταλόγου σε αποθήκευση με ευρετηρίαση μπορεί να είναι καταστροφική (τα μπλοκ δεδομένων δε «γνωρίζονται»).
- Στο UNIX κάθε αλλαγή που οδηγεί σε ανάθεση χώρου ή τροποποίηση των μεταδεδομένων γίνεται σύγχρονα, πριν γραφούν τα αντίστοιχα μπλοκ δεδομένων.
- Πρόβλημα: διακοπή σύγχρονης εγγραφής από πτώση τάσης.



Έλεγχος Συνέπειας σε UNIX (file system check - fsck)

```
***** FILE SYSTEM STILL DIRTY *****
** /dev/da0s1g
** Last Mounted on /external/Toshiba4
** Phase 1 - Check Blocks and Sizes
** Phase 2 - Check Pathnames
** Phase 3 - Check Connectivity
** Phase 4 - Check Reference Counts
LINK COUNT FILE I=423940 OWNER=bigbrother MODE=100644
SIZE=51 MTIME=Jan 24 17:55 2006 COUNT 2 SHOULD BE 1
ADJUST? [yn] █

UNREF FILE I=25062498 OWNER=root MODE=100600
SIZE=1013898 MTIME=Jan  8 08:04 2011
CLEAR? [yn] y

** Phase 5 - Check Cyl groups
FREE BLK COUNT(S) WRONG IN SUPERBLK
SALVAGE? [yn] y

SUMMARY INFORMATION BAD
SALVAGE? [yn] y

BLK(S) MISSING IN BIT MAPS
SALVAGE? [yn] y

184906 files, 30075175 used, 87991302 free (17902 frags, 10996675 blocks, 0.0% f
ragmentation)

***** FILE SYSTEM MARKED CLEAN *****

***** FILE SYSTEM WAS MODIFIED *****
```

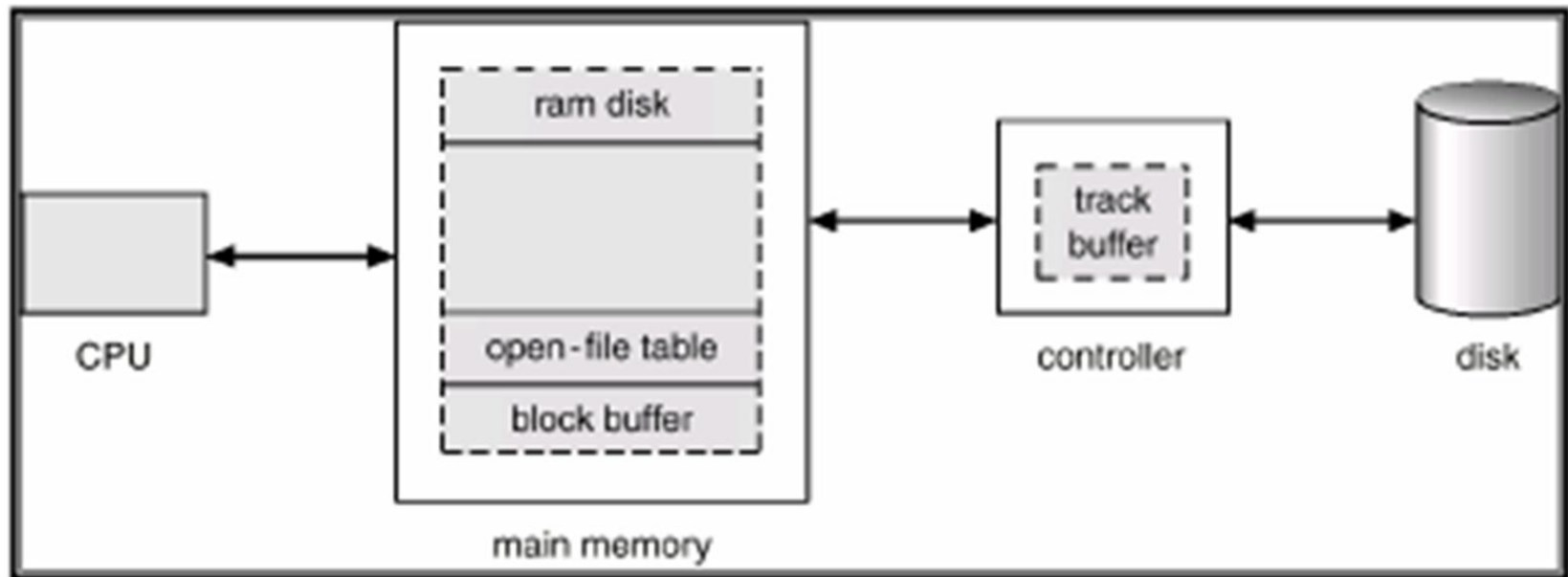


Βελτίωση της απόδοσης συστήματος αρχείων (1/2)

- Κρυφή μνήμη δίσκου (disk cache) στην κύρια μνήμη για την αποθήκευση των μπλοκ στα οποία γίνεται συχνή πρόσβαση.
- Χρήση εικονικού δίσκου (RAM disk) στην κύρια μνήμη.
- Υποστήριξη διαφορετικών μεθόδων πρόσβασης ανά κατηγορία εφαρμογών που χρησιμοποιούν το δίσκο.
- Έξυπνη υλοποίηση συστήματος αρχείων.
- Εγγραφή δεδομένων σε διαδοχικά μπλοκ/τομείς.
- Πρόωρη ανάγνωση μπλοκ ενός αρχείου.
- Καθυστερημένη εγγραφή μπλοκ των οποίων τα περιεχόμενα έχουν αλλάξει.



Βελτίωση της απόδοσης συστήματος αρχείων (2/2)



Συστήματα Αρχείων με Ημερολόγιο (Log/Journal) (1/2)

- Κάθε πρόσβαση στο σύστημα αρχείων καταγράφεται σε ημερολόγιο.
- Κάθε αλλαγή στο σύστημα αρχείων θεωρείται ότι έχει γίνει μόνο από τη στιγμή που έχει περαστεί στο ημερολόγιο (όχι πιο πριν) ανεξάρτητα από το αν το σύστημα αρχείων έχει ενημερώσει το δίσκο.
- Τα περιεχόμενα του ημερολογίου γράφονται ανεξάρτητα από τις λειτουργίες του συστήματος αρχείων σε σταθερή αποθήκευση (stable storage).



Συστήματα Αρχείων με Ημερολόγιο (Log/Journal) (2/2)

- Όταν ολοκληρωθεί μια αλλαγή στο σύστημα αρχείων, τότε σβήνεται η αντίστοιχη εισαγωγή από το ημερολόγιο.
- Αν το σύστημα αρχείων «πέσει», όταν το σύστημα επανέλθει, ελέγχεται το ημερολόγιο και εκτελούνται όλες οι λειτουργίες που εκκρεμούν.
- Αυτή η διαδικασία είναι πάρα πολύ γρήγορη (π.χ. αν δεν υπάρχει journal μπορεί να απαιτηθούν 30 λεπτά, αν υπάρχει journal μπορεί να απαιτηθούν 5 δευτερόλεπτα).

```
** [SU+] Recovering /dev/ad0s1h  
** Reading 33554432 byte journal from inode 166.  
** Building recovery table.  
** Resolving unreferenced inode list.  
** Processing journal entries.  
  
***** FILE SYSTEM MARKED CLEAN *****
```

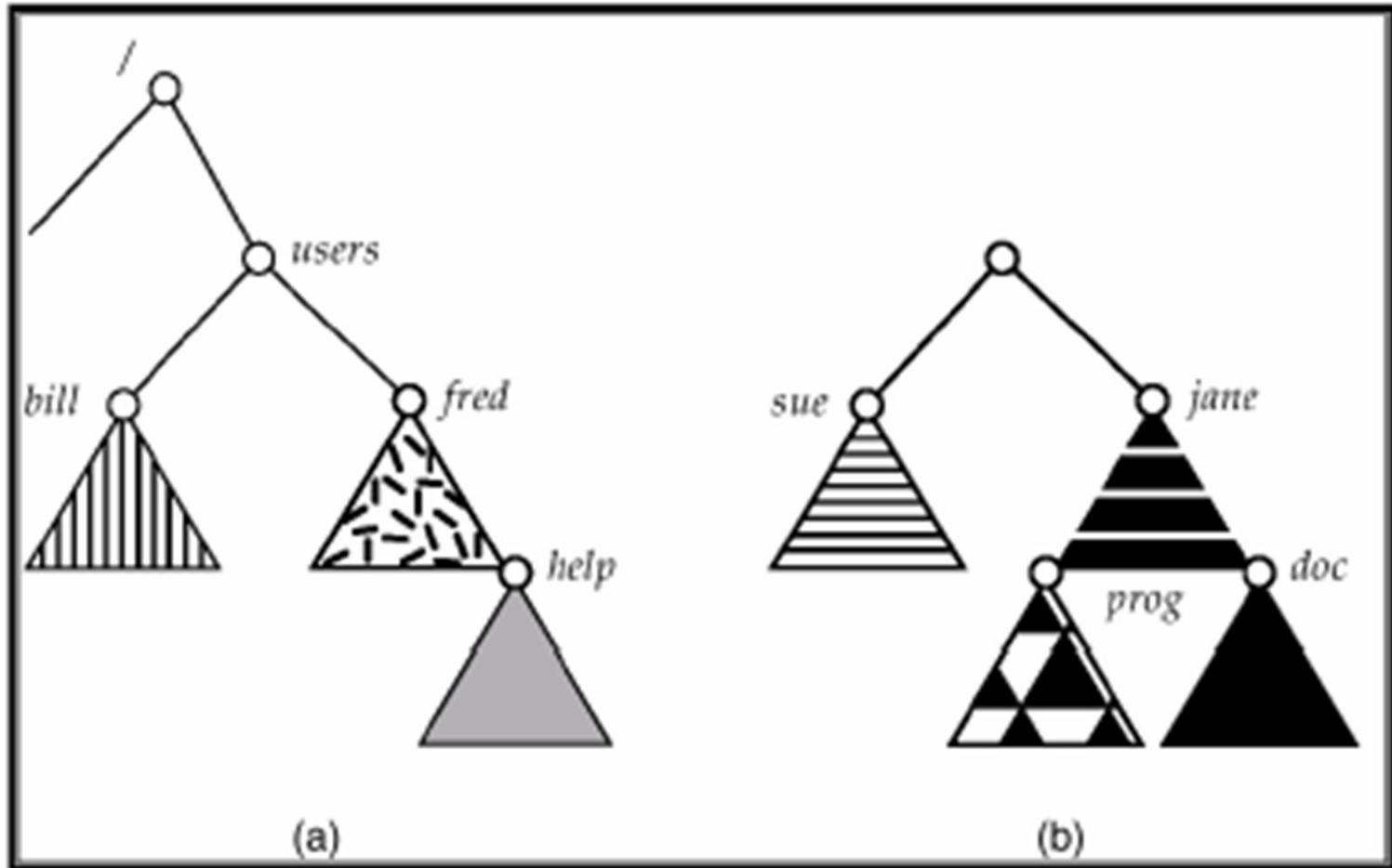


Προσάρτηση Συστήματος Αρχείων

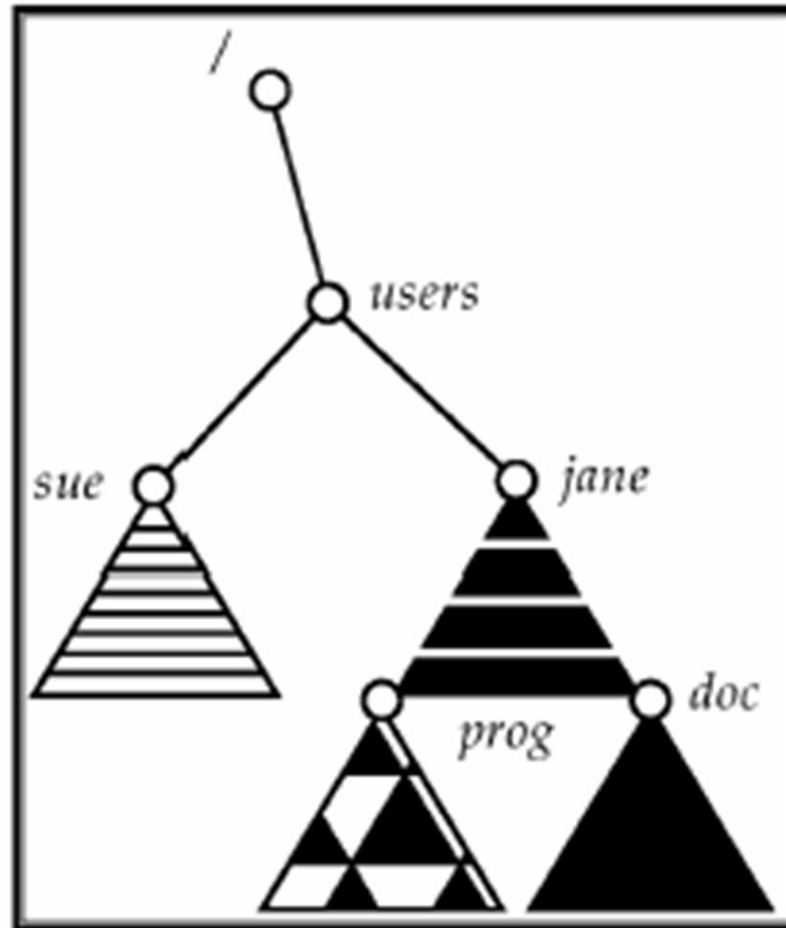
- Για να χρησιμοποιηθεί ένα σύστημα αρχείων πρέπει να προσαρτηθεί στο τοπικό σύστημα αρχείων (mounting a file system).
- Η προσάρτηση γίνεται σε συγκεκριμένο κατάλογο της ιεραρχίας του συστήματος, που προσδιορίζει ο διαχειριστής.
- Τα περιεχόμενα του καταλόγου αντικαθίστανται από τα περιεχόμενα του συστήματος που προσαρτήθηκε.



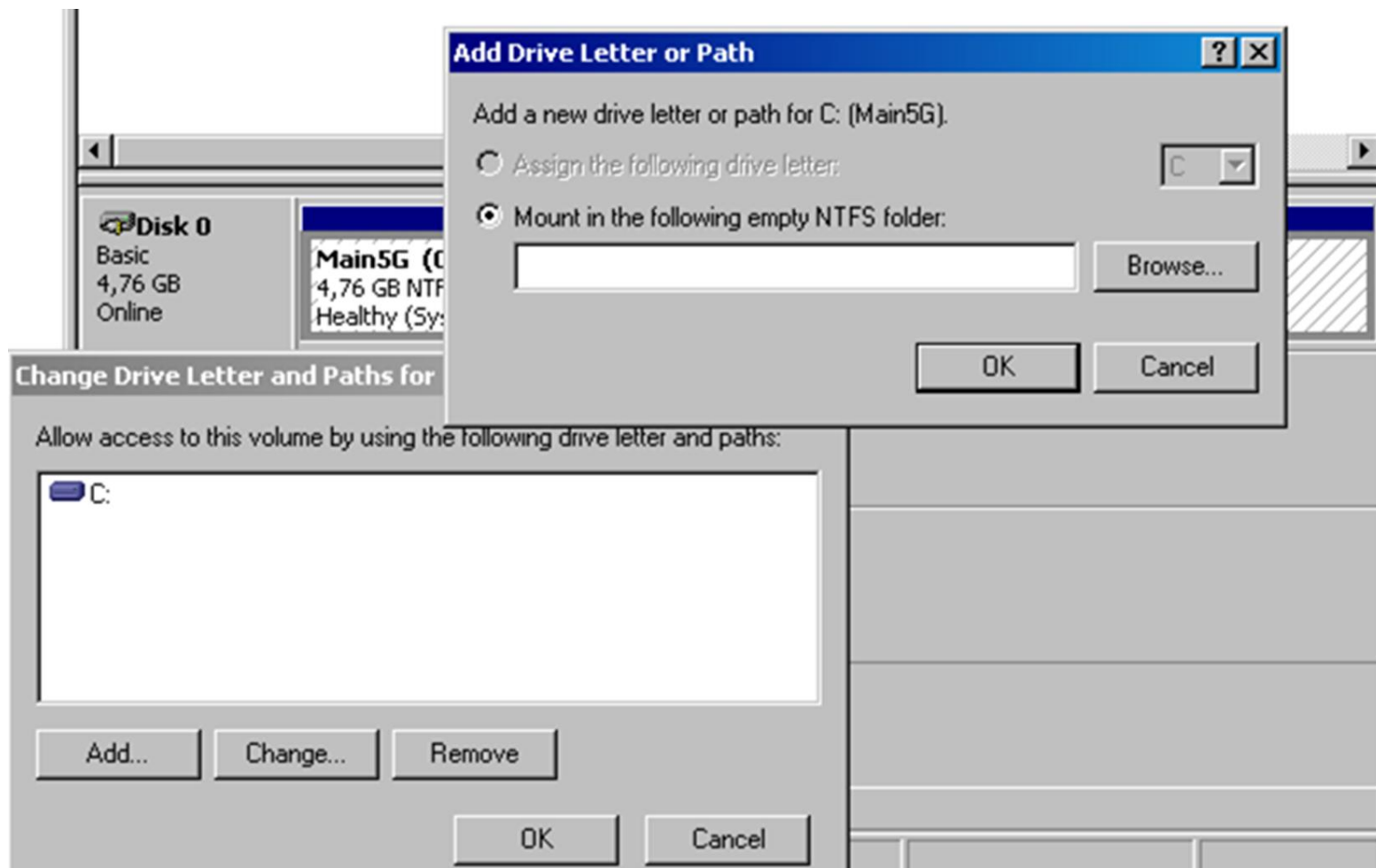
Πριν την προσάρτηση



Μετά την προσάρτηση



Προσάρτηση στα Windows XP



Προσάρτηση στο UNIX

```
therion# ls -l /external/Toshiba3/
total 0
therion# mount /dev/da0s1g /external/Toshiba4
therion# ls -la /external/Toshiba4
total 10
drwxr-xr-x   5 root   wheel   512 Apr 22 22:35 .
drwxr-xr-x   6 root   wheel   512 Apr 22 22:34 ..
drwxr-xr-x  25 root   wheel   512 Apr 30 10:25 .backupdir
drwxrwxr-x   2 root   operator 512 Oct  3  2009 .snap
drwxr-xr-x   4 root   wheel   512 Aug  1  2010 freebsd-update
therion# █
```

Υποστηριζόμενα συστήματα αρχείων για προσάρτηση στο FreeBSD 8

```
bigbrother@bigb5:[ ./bigbrother ]{104}> mount
mount          mount_mfs      mount_newnfs   mount_ntfs     mount_nwfs     mount_smbfs    mount_unionfs
mount_cd9660   mount_msdosfs mount_nfs       mount_nullfs   mount_portalfs mount_udf       mountd
ext2fs.ko     linprocfs.ko  msdosfs.ko     portalfs.ko    reiserfs.ko    ufs.ko         xfs.ko
fdescfs.ko    linsysfs.ko  ntfs.ko        procfs.ko      smbfs.ko        unionfs.ko     zfs.ko
geom_vol_ffs.ko mqueuefs.ko  nullfs.ko      pseudofs.ko    tmpfs.ko       usfs.ko
```



Προσάρτηση στο UNIX (εντολή mount)

```
/dev/sda5 on / type xfs (rw,noatime)
tmpfs on /lib/init/rw type tmpfs
  (rw,nosuid,mode=0755)
proc on /proc type proc
  (rw,noexec,nosuid,nodev)
sysfs on /sys type sysfs
  (rw,noexec,nosuid,nodev)
procbususb on /proc/bus/usb type usbfs (rw)
udev on /dev type tmpfs (rw,mode=0755)
tmpfs on /dev/shm type tmpfs (rw,nosuid,nodev)
  devpts on /dev/pts type devpts
  (rw,noexec,nosuid,gid=5,mode=620)
/dev/sda1 on /boot type ext3 (rw,noatime)
```



Περιορισμοί Χώρου

- Κάποια συστήματα αρχείων (ιδιαίτερα σε συστήματα πολλαπλών χρηστών), διαθέτουν την ικανότητα να ορίσει ο διαχειριστής μέγιστο αριθμό αρχείων και μπλοκ, ώστε να διασφαλιστεί ότι όλοι οι χρήστες θα έχουν δίκαια χρήση του αποθηκευτικού μέσου.
- FreeBSD (zfs):

```
[root@zafora /zstorage/home/mdasyg]# zfs get userquota@mdasygenis zstorage/home
NAME          PROPERTY          VALUE          SOURCE
zstorage/home userquota@mdasygenis 500M          local
[root@zafora /zstorage/home/mdasyg]# zfs get userused@mdasygenis zstorage/home
NAME          PROPERTY          VALUE          SOURCE
zstorage/home userused@mdasygenis 81.0M         local
[root@zafora /zstorage/home/mdasyg]#
```



Περιορισμοί Χώρου (Quota @ Windows)

The image displays three overlapping windows from the Windows operating system related to disk quotas on the Local Disk (C:).

Quota Entries for (C:)

Status	Name	Logon Name	Amount Used	Quota Limit	Warning Level	Percent Used
Warning	user	WM2007B5\user	1,07 GB	1,2 GB	1 GB	89
Warning	dmixalo	WM2007B5\dmixalo	1,19 GB	1,2 GB	1 GB	99
OK	[Retrieving Name]	S-1-5-32-544	7,39 GB	No Limit	No Limit	N/A

Quota Settings for (C:)

Quota

Status: Disk quota system is active

Enable quota management

Deny disk space to users exceeding quota limit

Select the default quota limit for new users on this volume:

Do not limit disk usage

Limit disk space to

Set warning level to

Select the quota logging options for this volume:

Log event when a user exceeds their quota limit

Log event when a user exceeds their warning level

Local Disk (C:) Properties

General Quota Customization

Type: Local Disk

File system: NTFS

Used space:	1,190,739,968 bytes	1,1C GE
Free space:	97,746,344 bytes	93,2 ME
Capacity:	1,288,486,312 bytes	1,1E GE

Drive C:

Compress this drive to save disk space

Allow files on this drive to have contents indexed in addition to file properties



Τι είναι τα i-nodes;

- Κάθε αρχείο σε σύστημα αρχείων στο UNIX έχει ένα μοναδικό περιγραφέα (ακέραιο αριθμό) που ονομάζεται i-node.
- Κάθε i-node συνήθως αντιστοιχεί σε block 8KB.
- Αν το αρχείο είναι μεγαλύτερο από 8KB τότε υπάρχει ο δείκτης προς το επόμενο i-nod.
- Αν δώσουμε την παράμετρο -i τότε μας εμφανίζονται στο ls τα i-nodes.

```
bigbrother@bigb5:[ ./bigbrother ]{342}> ls -li /etc/passwd  
50359 -rw-r--r-- 1 root wheel 2106 Mar 13 12:44 /etc/passwd
```

- Το i-node έχει τις πληροφορίες μετα-δεδομένων για το συγκεκριμένο block.

```
bigb5# find /etc -inum 50359  
/etc/passwd
```



UFS (Unix File System) (1/2)

- Χρησιμοποιείται από πλήθος UNIX based OS (MAC, Playstation 3).
- Ονομάζεται FFS (fast file system) στο BSD.
- Στην αρχή του δίσκου υπάρχουν:
 - τα boot blocks.
 - Το superblock (magic code, geometry, statistics).
 - Τοποθεσίες Cylinder groups.
- Τα cylinder groups (CG):
 - Έχουν αντίγραφο του superblock.
 - Header, free+used inodes, data blocks.
- Τα Directories έχουν μόνο τη λίστα των αρχείων και τα inodes που βρίσκεται το κάθε αρχείο.



UFS (Unix File System) (2/2)

- Χρησιμοποιούνται ομάδες CG λόγω του μεγάλου μεγέθους του δίσκου (καλύτερη τοπικότητα). Αν σε κάθε πρόσβαση πήγαινε στην αρχή του δίσκου η κεφαλή τότε θα υπήρχε μεγάλη καθυστέρηση.
- Χρησιμοποιείται block size 4KB.
- Χρησιμοποιούνται soft updates.
- Στο UFS2 υποστηρίζονται 64bit pointers (max capacity 8 zettabytes).
- Υποστηρίζονται soft-updates και snapshots.
- Υποστηρίζονται πολλαπλά partitions.
- Χρησιμοποιούνται fragments 2KB.

Multiples of bytes <small>v · d · e</small>				
SI decimal prefixes		Binary usage	IEC binary prefixes	
Name (Symbol)	Value		Name (Symbol)	Value
kilobyte (kB)	10^3	2^{10}	kibibyte (KiB)	2^{10}
megabyte (MB)	10^6	2^{20}	mebibyte (MiB)	2^{20}
gigabyte (GB)	10^9	2^{30}	gibibyte (GiB)	2^{30}
terabyte (TB)	10^{12}	2^{40}	tebibyte (TiB)	2^{40}
petabyte (PB)	10^{15}	2^{50}	pebibyte (PiB)	2^{50}
exabyte (EB)	10^{18}	2^{60}	exbibyte (EiB)	2^{60}
zettabyte (ZB)	10^{21}	2^{70}	zebibyte (ZiB)	2^{70}
yottabyte (YB)	10^{24}	2^{80}	yobibyte (YiB)	2^{80}

See also: [Multiples of bits](#) · [Orders of magnitude of data](#)



Δημιουργία συστήματος αρχείων

- Δημιουργία ενός νέου συστήματος αρχείων σε UFS που έχει ήδη τμηματοποιηθεί:

```
# newfs /dev/rsvnd0a
Warning: cylinder groups must have a multiple of 2 cylinders
Warning: 64 sector(s) in last cylinder unallocated
/dev/rsvnd0a: 65536 sectors in 656 cylinders of 1 tracks, 100 sectors
          32.0MB in 3 cyl groups (250 c/g, 12.21MB/g, 3136 i/g)
super-block backups (for fsck -b #) at:
  32, 25032, 50032,
```

- Παρατηρούμε ότι τοποθετούνται αντίγραφα του superblock σε συγκεκριμένες διευθύνσεις δίσκου.



Πληροφορίες για το σύστημα αρχείων UFS

- Χρησιμοποιούμε το `dumpfs`:

```
magic 19540119 (UFS2) time Sun May 22 06:15:04 2011
superblock location 65536 id [ 4c29e6b4 c56e3754 ]
ncg 4 size 262144 blocks 253815
bsize 16384 shift 14 mask 0xffffc000
fsize 2048 shift 11 mask 0xffff8000
frag 8 shift 3 fsbtodb 2
minfree 8% optim time symlinklen 120
maxbsize 16384 maxbpg 2048 maxcontig 8 contigsumsize 8
nbfree 19924 ndir 267 nifree 62038 nffree 1343
bpg 8193 fpg 65544 ipg 16448 unrefs 0
nindir 2048 inopb 64 maxfilesize 140806241583103
bsize 2048 cgsz 12288 csaddr 2112 cssize 2048
sblkno 40 cblkno 48 iblkno 56 dblkno 2112
cgrotor 0 fmod 0 ronly 0 clean 1
avgfdir 64 avgfilesize 16384
flags none
fsmnt /
volname swuid 0
```

```
cg 0:
magic 90255 tell 18000 time Sun May 22 03:51:22 2011
ncgx 0 ndblk 65544 niblk 16448 initiblk 3584 unrefs 0
nbfree 7117 ndir 31 nifree 16297 nffree 432
crotor 4320 irotor 109 frotor 4320
frsum 1 5 3 14 14 8 34
sum of frsum: 432
clusters 1-7: 7 6 4 7 2 1 3
clusters size 8 and over: 28
clusters free: 288, 298, 332, 340-344, 353-356, 391,
473-475, 536-542, 549-550, 560-574, 583-586, 601-661
```

```
3532-3535, 3630-4155, 4157-4158, 4160-8192
inodes used: 0-108, 688-691, 693-701, 703-719, 771-773, 775-782,
785
blks free: 2129-2135, 2154-2159, 2163-2167, 2171-2175, 2181-2183, 2203-2207,
2212-2215, 2220-2223, 2228-2235, 2240-2243, 2251-2255, 2260-2263,
2272-2278, 2284-2287, 2304-2318, 2320-2324, 2328-2334, 2340-2343,
2347-2351, 2355-2359, 2363-2367, 2371-2375, 2380-2391, 2440-2444,
2450-2455, 2466-2471, 2478, 2492, 2501, 2509, 2511, 2520, 2525
```



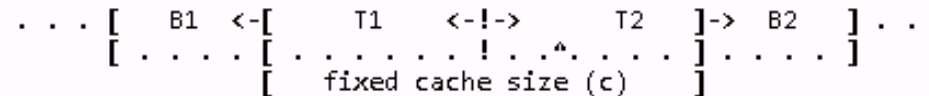
Το σύστημα αρχείων ZFS (ZetaByte File System)

- Δημιουργήθηκε από τη SUN (128bit file system).
- Είναι σύστημα αρχείων και διαχειριστής λογικών μονάδων (logical volume manager).
- Υποστηρίζει snapshots, data integrity check, automatic repair, quotas, deadline scheduling, redundancy, προστασία από silent corruption (SC: bit rot, cosmic radiation, current spikes, bugs in disk firmware, ghost writes),..
- Κανένα άλλο σύστημα αρχείων δεν παρέχει προστασία από SC (1 bit κάθε 10^{16}).
- Χρησιμοποιείται SHA-2 hash για κάθε block.
- Χρησιμοποιείται copy-on-write (σε νέο block).
- Δεν υπάρχει defragmentation (ακόμη).



Το ZFS χρησιμοποιεί την ARC (adaptive replacement cache)

- Αναπτύχθηκε το 2004 από την IBM.



- Χρησιμοποιεί 4 λίστες:
 - Πρόσφατες καταχωρήσεις (T1)
 - Πρόσφατες επαναλήψεις καταχωρήσεων (T2).
 - Καταχωρήσεις για απόσυρση από T1 (B1).
 - Καταχωρήσεις για απόσυρση από T2 (B2).
- Λειτουργία:
 - Αρχικά ένα στοιχείο μπαίνει στο T1 ωθώντας ένα στοιχείο προς το B1 (και τη διαγραφή κάποιου από B1).
 - Αν επαναληφθεί αναφορά σε αυτό πάει στο T2.
 - Αναλόγως αν υπάρχουν hit/mis T1,T2,B1,B2 τροποποιούνται τα μεγέθη της κάθε λίστας.



Το σύστημα αρχείων NTFS (1/2)

- Το default σύστημα αρχείων των windows (από 2000).
- Δημιουργήθηκε το 1994 (χρησιμοποιήθηκαν τμήματα του HPFS του OS/2).
- Είναι journaling file system.
- Hard-links, alternate data streams
"filename:streamname", sparse large files, compression, volume shadow copy, encrypting file system, quotas, transactional NTFS (ομάδα εντολών που ή θα εκτελεστούν όλες ή καμία), mount points, directory junctions (~symbolic links), timestamps in UTC.



Το σύστημα αρχείων NTFS (2/2)

- Όλα τα μεταδεδομένα αποθηκεύονται στην αρχή στην περιοχή \$MFT.
- 16-bit ονόματα αρχείων.
- Τα δεδομένα του συστήματος αρχείων αποθηκεύονται ως αρχεία στο ίδιο σύστημα αρχείων στον κατάλογο **“System Volume Information”**.
- 4KB block size.

```
C:\>dir /a
Volume in drive C has no label.
Volume Serial Number is CC33-D599

Directory of C:\

22/04/2011  06:39  μμ      <DIR>          $Recycle.Bin
14/07/2009  08:08  τμ      <JUNCTION>    Documents and Settings [C:\Users]
15/02/2012  07:03  μμ      <DIR>          emu8086
14/12/2010  12:18  μμ      <DIR>          Lexmark
24/08/2010  06:46  μμ      <DIR>          MSOCache
18/06/2013  12:07  τμ      1.073.741.824 pagefile.sys
14/07/2009  06:20  τμ      <DIR>          PerfLogs
17/03/2013  06:44  μμ      <DIR>          Program Files
17/03/2013  06:44  μμ      <DIR>          Program Files (x86)
18/02/2013  12:12  μμ      <DIR>          ProgramData
20/08/2010  04:52  μμ      <DIR>          Recovery
16/06/2013  09:22  τμ      <DIR>          System Volume Information
13/05/2013  02:16  μμ      <DIR>          Temp
22/04/2011  06:39  μμ      <DIR>          Users
26/02/2013  08:25  μμ      <DIR>          Windows
```



Προχωρημένα θέματα σε συστήματα αρχείων

- VFS (Virtual file system).
- RAID (redundant array of independent disks).
- NFS (network file system).
- SAN (storage area network).
- NAS (network attach storage), (A NAS is a single storage device that operate on data files, while a SAN is a local network of multiple devices that operate on disk blocks.)
- Clustering File Systems.
- FUSE (file system in userspace).
- Snapshots/Backup/Restore.



Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

