

Κεφάλαιο 2 - Οργάνωση Συστήματος Αρχείων

Αρχείο (File)

Αρχείο ονομάζεται μία οργανωμένη συλλογή από δεδομένα που είναι αποθηκευμένα σε κάποια συσκευή αποθήκευσης στον υπολογιστή.

Για να κρατήσουμε μία εργασία που κάνουμε στον Η/Υ (π.χ. ένα κείμενο, μία ζωγραφιά) την αποθηκεύουμε στο σκληρό δίσκο του Η/Υ αφού πρώτα της δώσουμε κάποιο περιγραφικό όνομα. Μια εργασία την οποία έχουμε αποθηκεύσει στο σκληρό δίσκο, ονομάζεται **αρχείο δεδομένων (Data File)**. Ένα αρχείο δεδομένων εμφανίζεται από το windows ως ένα εικονίδιο με το όνομα που έχουμε δώσει. Με τον ίδιο τρόπο ένα σύνολο από εντολές (δηλαδή ένα πρόγραμμα, το οποίο το δημιούργησε κάποιος ειδικός στους Η/Υ - ένας προγραμματιστής) που βρίσκεται αποθηκευμένο στο σκληρό δίσκο ονομάζεται **αρχείο προγράμματος(Program File)**.

Ευρετήριο (Directory) ή Φάκελος (Folder) ή Κατάλογος (Catalogue)

Όπως στο σπίτι μας τακτοποιούμε τα χαρτιά μας σε κατηγορίες και τα τοποθετούμε σε φακέλους και ντοσιέ, έτσι ταξινομούμε και τα αρχεία μας στον σκληρό δίσκο του Η/Υ. Χωρίζουμε το σκληρό δίσκο σε νοητές περιοχές για να τοποθετήσουμε τα αρχεία μας ταξινομημένα σε θεματικές κατηγορίες.

Οι ξεχωριστές περιοχές που δημιουργούμε – στις οποίες και δίνουμε κάποιο περιγραφικό όνομα - στο σκληρό δίσκο του Η/Υ και στις οποίες τοποθετούμε έναν αριθμό αρχείων με σκοπό να διατηρήσουμε τα αρχεία μας τακτοποιημένα είναι ειδικά αρχεία ενός ΛΣ τα οποία ονομάζονται **Ευρετήρια**. Αυτά εμφανίζονται από τα windows ως εικονίδια με κίτρινο χρώμα και σχήμα καρτέλας.

Οργάνωση Αρχείων

Σύστημα αρχείων ονομάζουμε τον τρόπο με τον οποίο, ένα ΛΣ οργανώνει τα αρχεία. Για την οργάνωση των αρχείων μας δημιουργούμε έναν αριθμό από ευρετήρια χρησιμοποιώντας τη **δενδροειδή δομή**, όπως ακριβώς κάνουμε με το γενεαλογικό δένδρο.

Με τον τρόπο αυτό προκύπτει ένα δένδρο όπου τα κλαδιά του αντιπροσωπεύουν ευρετήρια και τα φύλλα αντιπροσωπεύουν αρχεία.

Το σημείο από το οποίο ξεκινάμε να δημιουργούμε ευρετήρια ονομάζεται **ρίζα** του δένδρου μας. Η ρίζα σε έναν Η/Υ είναι η αντίστοιχη συσκευή μόνιμης αποθήκευσης στην οποία θέλουμε να αποθηκεύσουμε τα αρχεία μας.

Τα ευρετήρια που δημιουργούμε μέσα σε ένα ευρετήριο ονομάζονται **υποευρετήρια** του, ενώ το ευρετήριο μέσα στο οποίο έχουμε δημιουργήσει κάποια άλλα ευρετήρια ονομάζεται **γονέας** τους.

Το ευρετήριο στο οποίο βρίσκεται ο χρήστης κάθε φορά λέγεται **τρέχον ή ευρετήριο εργασίας**.

Δύο αρχεία (ή ευρετήρια) επιτρέπεται να έχουν το ίδιο όνομα, αρκεί να έχουν διαφορετικό γονέα.

Στα windows οι συσκευές μόνιμης αποθήκευσης συμβολίζονται με τα παρακάτω λατινικά γράμματα :

A: ο μηχανισμός δισκέτας

C: ο σκληρός δίσκος

D: η μονάδα CD-ROM (ή ο δεύτερος σκληρός δίσκος αν υπάρχει)

Απόλυτο όνομα (Absolute pathname) ενός αρχείου ή μονοπάτι (path) ενός αρχείου ονομάζομαι τα ευρετήρια, με τη σειρά που πρέπει να τα προσπελάσουμε (ανοίξουμε) ξεκινώντας από τη ρίζα, μέχρι να βρούμε το αρχείο που μας ενδιαφέρει. Ανάμεσα στα ευρετήρια αυτά βάζουμε το σύμβολο κάθετος (\ ή /) π.χ **C:\Τα έγγραφα μου\κείμενα\εργασίες\υγιεινή διατροφή.doc**

Σχετικό όνομα ενός αρχείου ονομάζομαι τα ευρετήρια, με τη σειρά που πρέπει να τα προσπελάσουμε (ανοίξουμε) ξεκινώντας από το τρέχον ευρετήριο, μέχρι να βρούμε το αρχείο που μας ενδιαφέρει.

..\εργασίες\υγιεινή διατροφή.doc

Ένα αρχείο προσδιορίζεται όχι μόνο από το όνομα του αλλά και από το μονοπάτι, τη θέση του δηλαδή μέσα στο σκληρό δίσκο.

Ονοματολογία

Το όνομα ενός ευρετηρίου ή αρχείου μπορεί να έχει μήκος ως 255 χαρακτήρες (επιτρέπονται και τα κενά) και απαγορεύεται να περιλαμβάνει του χαρακτήρες * | \ / > < ? ' ' .

Χαρακτηριστικά

Κάθε αρχείο έχει κάποια χαρακτηριστικά για λόγους προστασίας όπως είναι **Μόνο για Ανάγνωση (Read Only)**, **Προς Φύλαξη (Archive)**, **Κρυφό (Hidden)** και **Συστήματος (System)**.

Δικαιώματα

Κάθε χρήστης έχει κάποια δικαιώματα χρήσης κάθε αρχείου ξεχωριστά όπως είναι η **Εγγραφή (Write)**, η **Ανάγνωση (Read)**, η **Τροποποίηση (Modify)** και η **Εκτέλεση (Run)**

Κατηγορίες αρχείων

Εκτός από τα **αρχεία Δεδομένων** και **Προγράμματος** υπάρχουν επίσης :

Αρχεία Κειμένου (Text Files)

Αρχεία Συστήματος (System Files)

Αρχεία Συσκευών (Device Files)

Προσωρινά Αρχεία (Temporary Files)

Αρχεία Εκτύπωσης (Printer, Spooler Files)

Εφεδρικά Αρχεία (Backup Files)

Αρχεία Δέσμης Εντολών (Batch Files)

Τι είναι η επέκταση ενός αρχείου ;

Σε κάθε αρχείο, όπως ξέρουμε δίνουμε ένα συγκεκριμένο όνομα - το οποίο φροντίζουμε να είναι αρκετά περιγραφικό - και το αποθηκεύουμε σε συγκεκριμένη θέση στο σκληρό δίσκο του Η/Υ μας. Μετά το όνομα που έχουμε δώσει εμείς, η εφαρμογή με την οποία το δημιουργήσαμε προσθέτει μία τελεία και μία επέκταση από 3 ή 4 λατινικούς χαρακτήρες, η οποία δηλώνει το είδος του αρχείου (από ποια εφαρμογή δημιουργήθηκε).

Οι πιο συνηθισμένες επεκτάσεις αρχείων είναι:

.txt, .rtf	αρχείο με απλό κείμενο
.doc, .docx	αρχείο με κείμενο του word
.pdf	αρχείο με κείμενο του acrobat reader
.xls, .xlsx	αρχείο του excel
.bmp, .jpg, .gif, .tif, .png	αρχείο με εικόνα
.wav, .mp3, .wma, .m3u, .mid	αρχείο με ήχο
.avi, .mov, .mpg, .swf, .flv, .mp4, .3pg	αρχείο με ταινία
.htm, .html, .php	αρχείο ιστοσελίδας

Τι είναι ο Διαχειριστής Αρχείων (File Manager) ;

Διαχειριστή Αρχείων ονομάζουμε το πρόγραμμα που χρησιμοποιείται για την διαχείριση των φακέλων και των αρχείων σε έναν υπολογιστή. Στα windows θα το βρείτε με το όνομα "Εξερεύνηση των Windows".

Ποιες εργασίες μπορούμε να εκτελέσουμε σε ένα αρχείο ή φάκελο ;

1. Δημιουργία για ευρετήρια
2. Αναζήτηση
3. Μετονομασία
4. Διαγραφή
5. Μετακίνηση - Αντιγραφή
6. Δημιουργία συντόμευσης
7. Συμπίεση
8. Αλλαγή χαρακτηριστικών και δικαιωμάτων

Φυσική οργάνωση δεδομένων

Πριν προχωρήσουμε στην μορφοποίηση (format) ενός δίσκου ώστε να μπορεί να χρησιμοποιηθεί για την αποθήκευση δεδομένων, πρέπει πρώτα να τον χωρίσουμε σε διαμερίσματα (partitions) ή τόμους (volumes).

2.1.9 Αποθήκευση δεδομένων

Το διαβάζετε από το βιβλίο

2.1.10 Προσπέλαση Δίσκων

Το διαβάζετε από το βιβλίο

Είδη συστημάτων αρχείων

Κατά τη διαμόρφωση ορίζουμε το σύστημα αρχείων που θα χρησιμοποιηθεί :

1. FAT (File Allocation Table) / FAT32

Πρωτοχρησιμοποιήθηκε στο DOS / windows 95, με μεγάλη υποστήριξη συσκευών.

Ενδείκνυται για αρχεία μικρού μεγέθους και συσκευές μικρής χωρητικότητας.

Υποστηρίζει μέγεθος δίσκου ως 32GB και επιτρέπει τη διαχείριση αρχείων μεγέθους ως 2 / 4 GB.

2. NTFS (New Technology File System)

Πρωτοχρησιμοποιήθηκε στα windows NT

Υποστηρίζει μέγεθος δίσκου ως 256TB και επιτρέπει τη διαχείριση αρχείων μεγέθους ως 16 TB.

3. exFAT (Extended FAT)

Σύστημα της Microsoft σχεδιασμένο για μνήμες flash, με περιορισμένη όμως υποστήριξη συσκευών.

4. Ext4 (Extended File System)

Χρησιμοποιείται στις εκδόσεις Linux.

Υποστηρίζει μέγεθος δίσκου ως 1 EB και επιτρέπει τη διαχείριση αρχείων μεγέθους ως 16 TB.

Στο NTFS υπάρχει η δυνατότητα να χωρίσουμε ένα δίσκο μέχρι και σε 3 **πρωτεύοντα διαμερίσματα** (εκκινήσιμα). Ο υπόλοιπος τυχόν αδιάθετος χώρος δηλώνεται ως 1 **εκτεταμένο διαμέρισμα** στο οποίο δημιουργούμε μία σειρά από **λογικούς τόμους**.

Κατά το χωρισμό σε διαμερίσματα σε κάθε εκκινήσιμο δίσκο δημιουργείται η εγγραφή MBR (Master Boot Record) στην οποία αποθηκεύεται ο Boot Manager.

Σήμερα οι συσκευές Δίσκου διαβάζουν μονομιάς μία σειρά από 2 ως 64 συνεχόμενους τομείς των 512 B. Μια τέτοια ομάδα τομέων ονομάζεται συστοιχία (cluster) και αποτελεί τη μικρότερη περιοχή του δίσκου που μπορεί να αποδοθεί σε ένα αρχείο. Είναι γνωστή ως **Μονάδα εκχώρησης (block)**. Αυτή έχουμε τη δυνατότητα (και στο NTFS) να την ορίσουμε κατά τη διαδικασία της μορφοποίησης.

Εκτός από βασικούς δίσκους στους οποίους κάθε αρχείο βρίσκεται σε ένα δίσκο, έχουμε και τους **δυναμικούς δίσκους** όπου ένα αρχείο μπορεί να μοιράζεται σε πάνω από 2 δίσκους (**συστοιχίες RAID**)

2.1.11 Καταχώρηση και απελευθέρωση περιοχών του δίσκου

Όταν φτάσει η στιγμή να δημιουργηθεί ένα αρχείο, το Λειτουργικό Σύστημα το αποθηκεύει σε **μονάδες εκχώρησης (συστοιχίες, clusters)**. Κάθε αρχείο έχει τουλάχιστον μια μονάδα εκχώρησης. Αυτό σημαίνει ότι, αν η μονάδα εκχώρησης έχει μέγεθος 4096 bytes και το αρχείο έχει περιεχόμενο έναν χαρακτήρα (δηλαδή 1 byte), τότε ο χώρος που θα καταλαμβάνει το αρχείο στον δίσκο θα είναι 4096 bytes. Τα υπόλοιπα 4095 bytes λοιπόν δεν θα αξιοποιούνται. Αυτό ονομάζεται **εσωτερικός κατακερματισμός (internal fragmentation)** του δίσκου. Το πρόβλημα που υπάρχει λοιπόν είναι ότι αν ορίσουμε μεγάλη τιμή στην μονάδα εκχώρησης τότε δεν θα αξιοποιηθεί σωστά ο χώρος στην περίπτωση που αποθηκεύουμε πολλά μικρά αρχεία.

Από την άλλη αν ορίσουμε μικρή τιμή στην μονάδα εκχώρησης τότε ένα μεγάλο αρχείο θα πρέπει να χωριστεί σε πάρα πολλά μικρά κομμάτια. Έτσι αν ως μονάδα εκχώρησης σε ένα δίσκο ορίσουμε την τιμή 1 KB και αποθηκεύσουμε σε αυτόν ένα αρχείο μεγέθους 2 GB τότε θα χρειαστούν για την αποθήκευση του 2 εκατομμύρια μονάδες εκχώρησης, γεγονός που θα έχει ως αποτέλεσμα μεγάλη καθυστέρηση στη προσπέλαση του αρχείου.

Η απόδοση των μονάδων εκχώρησης (blocks, μπλοκ) στα αρχεία γίνεται με διάφορους τρόπους οι οποίοι αναφέρονται παρακάτω :

- **Συνεχής καταχώρηση (Contiguous Allocation)**

Στην καταχώρηση αυτή τα μπλοκ του αρχείου είναι συνεχόμενα στον δίσκο. Αυτό έχει το πλεονέκτημα ότι είναι απλό στην υλοποίηση και ότι για κάθε αρχείο απαιτείται μόνο η διεύθυνση του πρώτου μπλοκ. Το μειονέκτημα είναι ότι το μέγεθος των αρχείων δεν είναι πάντα γνωστό κατά τη στιγμή της δημιουργίας τους και το σύστημα δεν γνωρίζει πόσο χώρο να δεσμεύσει.

- **Καταχώρηση συνδεδεμένης λίστας (Linked List Allocation)**

Σε αυτή τη μέθοδο το αρχείο καταχωρείται ως μια συνδεδεμένη λίστα από μπλοκ. Στο τέλος του πρώτου μπλοκ τοποθετείται ο αριθμός του επόμενου μπλοκ (ένας δείκτης δηλαδή στο επόμενο μπλοκ) και το τελευταίο μπλοκ έχει μια ειδική τιμή για να δείξει το τέλος της αλυσίδας. Ένα μειονέκτημα αυτής της μεθόδου είναι ότι δεν είναι δυνατή η άμεση προσπέλαση σε κάποιο τμήμα του αρχείου καθώς δεν είναι γνωστές οι διευθύνσεις των μπλοκ.

- **Καταχώρηση με χρήση δείκτη (Indexed Allocation)**

Αυτή είναι μια παραλλαγή της μεθόδου συνδεδεμένης λίστας στην οποία διατηρείται ένας πίνακας όπου υπάρχει μια θέση για κάθε μπλοκ του δίσκου (FAT, File Allocation Table). Έτσι, το περιεχόμενο της θέσης του πίνακα που αντιστοιχεί στο πρώτο μπλοκ του αρχείου θα είναι η διεύθυνση του δεύτερου μπλοκ του αρχείου κ.ο.κ. Στη θέση του τελευταίου μπλοκ υπάρχει επίσης μια ειδική τιμή που σηματοδοτεί το τέλος του αρχείου. Όπως φαίνεται και από το όνομα του πίνακα αυτή η μέθοδος χρησιμοποιείται στα συστήματα FAT που είδαμε παραπάνω.

- **Κόμβοι-δ (i-nodes)**

Η μέθοδος αυτή στηρίζεται πάλι στη χρήση δεικτών αλλά με διαφορετικό τρόπο. Έτσι, για κάθε αρχείο υπάρχει ένας μικρός πίνακας που λέγεται δ-κόμβος (i-node) και αυτός περιέχει τους αριθμούς των μπλοκ του αρχείου. Σε περίπτωση μεγάλων αρχείων όπου ο πίνακας δεν είναι αρκετός για να χωρέσει όλες τις θέσεις των μπλοκ, μια θέση του πίνακα αυτού περιέχει την διεύθυνση έναν άλλου πίνακα που περιέχει τις υπόλοιπες θέσεις. Σε αυτή τη λογική στηρίζονται τα συστήματα ext2, ext3, ext4. Επίσης το σύστημα NTFS υλοποιεί κάτι ανάλογο με τη χρήση του Master File Table (MFT).

2.1.12 Κατακερματισμός (Fragmentation).

Μετά από πολλές διαδικασίες δημιουργίας και διαγραφής αρχείων είναι αναμενόμενο ότι τα μπλοκ του κάθε αρχείου θα βρίσκονται διασκορπισμένα στον δίσκο. Αυτό θα έχει ως αποτέλεσμα οι φυσικές διαδικασίες ανάγνωσης/εγγραφής να απαιτούν πολλές μετακινήσεις των κεφαλών του σκληρού δίσκου έτσι ώστε να βρεθούν στις κατάλληλες θέσεις. Έτσι όμως αυξάνεται ο χρόνος που απαιτείται για την ανάγνωση/εγγραφή. Αυτή η κατάσταση ονομάζεται **εξωτερικός κατακερματισμός (external fragmentation)** και μειώνει την απόδοση του δίσκου. Για την αντιμετώπιση του είναι δυνατό να γίνει μια **λειτουργία ανασυγκρότησης (αποκατακερματισμού, defragmentation)** του δίσκου όπου τα μπλοκ των αρχείων τοποθετούνται σε γειτονικές θέσεις στο μεγαλύτερο δυνατό βαθμό. Η λειτουργία αυτή είναι καλό να γίνεται τακτικά σε κάθε δίσκο ενός υπολογιστικού συστήματος.

2.2 Ασφάλεια συστήματος

Το διαβάζετε από το βιβλίο